

# IMPLICIT MASS-MATRIX PENALIZATION OF HAMILTONIAN DYNAMICS WITH APPLICATION TO EXACT SAMPLING OF STIFF SYSTEMS

PETR PLECHÁČ\* AND MATHIAS ROUSSET†

**Abstract.** An implicit mass-matrix penalization (IMMP) of Hamiltonian dynamics is proposed, and associated dynamical integrators, as well as sampling Monte-Carlo schemes, are analyzed for systems with multiple time scales. The penalization is based on an extended Hamiltonian with artificial constraints associated with some selected DOFs. The penalty parameters enable arbitrary tuning of timescales for the selected DOFs. The IMMP dynamics is shown to be an interpolation between the exact Hamiltonian dynamics and the dynamics with rigid constraints. This property translates in the associated numerical integrator into a tunable trade-off between stability and dynamical modification. Moreover, a penalty that vanishes with the time-step yields order two convergent schemes for the exact dynamics. Moreover, by construction, the resulting dynamics preserves the canonical equilibrium distribution in position variables, up to a computable geometric correcting potential, leading to Metropolis-like unbiased sampling algorithms. The algorithms can be implemented with a simple modification of standard geometric integrators with algebraic constraints imposed on the selected DOFs, and has no additional complexity in terms of enforcing the constraints and force evaluations. The properties of the IMMP method are demonstrated numerically on the  $N$ -alkane model, showing that the time-step stability region of integrators and the sampling efficiency can be increased with a gain that grows with the size of the system. This feature is mathematically analyzed for a harmonic atomic chain model. When a large stiffness parameter is introduced, the IMMP method is shown to be asymptotically stable and to converge towards the heuristically expected Markovian effective dynamics on the slow manifold.

**Key words.** Hamiltonian systems, NVT ensemble, stiff dynamics, Langevin dynamics, constrained dynamics, Hybrid Monte Carlo.

**AMS subject classifications.** 65C05, 65C20, 82B20, 82B80, 82-08

**1. Introduction.** This paper deals with numerical integration and sampling of Hamiltonian systems with multiple timescales. The main motivation is to develop numerical integration methods for the dynamics which resolves certain selected fast degrees of freedom only "statistically", and that can be also used to sample accurately the canonical equilibrium distribution. Furthermore, as an ultimate goal, one also seeks good approximation of dynamical behavior, at least at large temporal scales.

Hamiltonian systems with multiple timescales typically appear in molecular dynamics (MD) simulations, which have become, with the aid of increasing computational power, a standard tool in many fields of physics, chemistry and biology. However, extending the simulations to physically relevant time-scales remains a major challenge for various large molecular systems. Due to the complexity of implicit methods, the time scales reachable by standard numerical methods are usually limited by the rapid oscillations of some particular degrees of freedom. Since the sampling dynamics has to be integrated for long times, the time-step restriction associated with fast oscillations/short time scales in molecular systems contributes to the high computational cost of such methods. However, the physical necessity of resolving the fast degrees of freedom in simulations is often ambiguous, and efficient treatment of the fast time scales has motivated new interest in developing numerical schemes for the integration of such stiff systems.

The problem of integrating stiff forces is relevant both for the direct numerical simulation of the Hamiltonian dynamics, as well as for the less restrictive problem of designing a sampling scheme with respect to the canonical ensemble. Sampling from the canonical distribution can be achieved by Markov chain Monte Carlo (MCMC) algorithms based on a priori knowledge of possible moves combined with a Metropolis-Hastings acceptance/rejection corrector (a historical reference is [31]). For complex molecular systems, however, such global moves remain unknown in general, and sampling methods consists generically in using either a Hamiltonian dynamics integrator with a thermostat (e.g., a Langevin process), or its overdamped limit, a drifted random walk (Brownian dynamics) (see [7] for a review and references on

---

\*Department of Mathematics, University of Tennessee, Knoxville, TN 37996-1300 and Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA, (plechac@math.utk.edu).

†INRIA Lille Nord-Europe, Villeneuve d'Ascq, France, (mathias.rousset@inria.fr)

classical sampling methods). Brownian dynamics of systems with multiple time scales suffers from similar stability restrictions (see [20, 41] for some practical issues related to Brownian dynamics simulations in MD).

Broadly speaking, one may start by recognizing two approaches to the numerical treatment of stiff systems:

- (i) Semi-implicit, multi-step integrators and their variants (e.g., the textbooks [36, 14], or the review paper [10] and references therein, [20] for Brownian dynamics), which attempt to resolve microscopic highly oscillatory dynamical behavior.
- (ii) Methods with direct constraints, where the highly oscillatory degrees of freedom are constrained to their equilibrium value (e.g., [36, 28, 40, 19] and references therein).

In spite of their differences the common key feature of all these methods is to balance a trade-off between stability restrictions and implicit time-stepping form, or in other words, between the computational effort associated with small time steps, and the computational cost of solving implicit equations implied by the stiffness.

Although constrained dynamics remove, in principle, the stiffness of the associated numerical scheme, it introduces new difficulties and numerical problems. As an approximation to the original dynamics it modifies important features of the system; most importantly, the original statistical distribution. The principal goal of the proposed method is to replace direct constraints by implicit mass-matrix penalization (IMMP), detailed in Section 3, which integrates fDOFs, but with a tunable mass penalty. The method designed in this way achieves the two goals:

- (i) from the dynamical point of view, the IMMP method amounts to an appropriate interpolation between exact dynamics and constrained dynamics considered in the second family of the methods mentioned above. Moreover, a freely tunable trade-off between dynamical modification and stability is obtained.
- (ii) from the sampling point of view, the IMMP dynamics preserves the canonical equilibrium distribution, up to a time step error and an easily computable geometric correcting potential. This leads to Metropolis Monte Carlo methods that sample *exactly* the canonical distribution. When using Metropolis schemes, the forces arising from the geometric correcting potential need not be computed.

The idea of adjusting mass tensors in order to slow down fast degrees of freedom goes back to [3]. In this paper, the author proposes to modify the mass tensor with respect to the Hessian of the potential energy function in order to confine the frequency spectrum to low frequencies only. Two natural drawbacks of this procedure arise from the costly computation of the second-order derivatives of the potential, and from the bias introduced when the adjusted mass-tensor is adapted during the dynamics. Such an approach seems inevitable when the fDOFs are unknown, but in many cases, the fast degrees of freedom are explicitly given by the structure of the system (e.g., co-valent and angle bonds in molecular chains). To our knowledge, mass tensor modification have been used in practical MD simulations by increasing the mass of some well-chosen (e.g., light) atoms [26, 29]. The aim of this paper is to propose a more systematic mass-tensor modification strategy.

The proposed method relies on the assumption that the system Hamiltonian is separable with quadratic kinetic energy

$$H(p, q) = \frac{1}{2} p^T M^{-1} p + V(q), \quad (1.1)$$

and that the “fast” degrees of freedom  $(\xi_1, \dots, \xi_n)$  are explicitly defined, smooth functions of the system position

$$q = (q_1, \dots, q_d) \mapsto (\xi_1(q), \dots, \xi_n(q)). \quad (1.2)$$

We emphasize that the knowledge of “fast forces” is not required, and the variables  $\xi$  can be chosen arbitrarily. If the fDOFs are not identified the method retains its approximation properties while not

performing efficiently. The fDOFs are penalized with a mass-tensor modification given by

$$M_\nu(q) = M + \nu^2 \nabla_q \xi M_z \nabla_q^T \xi, \quad (1.3)$$

where  $\nu$  denotes the penalty intensity, and  $M_z$  a “virtual” mass matrix associated with the fDOFs. The modification does not impact motions orthogonal to the fDOFs. The position dependence of the mass-penalization introduces a geometric bias. This bias is corrected by introducing an effective potential

$$V_{\text{fix},\nu}(q) = \frac{1}{2\beta} \ln(\det(M_\nu(q))), \quad (1.4)$$

which will turn out to be a  $\nu^{-1}$ -perturbation of the usual Fixman corrector (see [19]) associated with the sub-manifold defined by constraining the fDOFs  $\xi$ . The key point is then to use an implicit representation of the mass penalty with the aid of the extended Hamiltonian

$$\begin{cases} H_{\text{IMMP}}(p, p_z, q, z) = \frac{1}{2} p^T M^{-1} p + \frac{1}{2} p_z^T M_z^{-1} p_z + V(q) + V_{\text{fix},\nu}(q), \\ \xi(q) = \frac{z}{\nu}. \end{cases} \quad (C_\nu) \quad (1.5)$$

The auxiliary degrees of freedom  $z$  are endowed with the “virtual” mass-matrix  $M_z$ . The constraints  $(C_\nu)$  are applied in order to identify the auxiliary variables and the fDOFs  $\xi$  with a coupling intensity tuned by  $\nu$ . The typical time scale of the fDOFs is thus enforced by the penalty  $\nu$ . The system is coupled to a thermostat through a Langevin equation (2.3), which yields a stochastically perturbed dynamics that samples the equilibrium canonical distribution. We then obtain the following desirable properties:

1. The associated canonical equilibrium distribution in position is independent of the penalty  $\nu$ .
2. The limit of vanishing penalization ( $\nu = 0$ ) is the original full dynamics, enabling the construction of dynamically consistent numerical schemes.
3. The limit of infinite penalization is a standard effective constrained dynamics on the “slow” manifold associated with stiff constraints on  $\xi$ .
4. Numerical integrators can be obtained through a simple modification of standard integrators for effective dynamics with constraints yielding equivalent computational complexity.

The dynamics associated with the IMMP Hamiltonian (1.5) is detailed in (3.5), see Section 3. The numerical discretization (using a leapfrog/Verlet splitting with constraints, usually called “RATTLE” for fully constrained dynamics) is given by (4.1). When considering sampling, the time-step error of the numerical flow can be corrected with a Metropolis step (the so-called Generalized Hybrid Monte-Carlo method, see references in Section 4) to obtain *exact sampling*. When this correction is introduced, the gradient of the Fixman potential (1.4) need not be computed. These numerical aspects of the method are presented in Section 4. By including a penalty, the proposed method modifies the original Hamiltonian. However, the mass penalty can be also thought of as depending on the time step  $\nu = \nu(\delta t)$  leading to order two consistent schemes.

In Section 6, we introduce a small stiffness parameter  $\epsilon$  encoding the fastest DOFs, and show that the penalty intensity can be scaled as  $\nu = \bar{\nu}/\epsilon$  in order to obtain asymptotically stable dynamics in the limit  $\epsilon \rightarrow 0$ . We prove that the dynamics converge towards the expected Markovian effective dynamics on the slow manifold. We also present analysis of the corresponding asymptotic preserving properties of the proposed scheme.

High-dimensional systems usually contain a large variety of timescales, and are therefore challenging test cases. The  $N$ -alkane model is numerically studied in Section 5 and systematically compared to Verlet scheme and constrained integration, with separate studies for dynamical and sampling issues. In the case of butane, bond angles are penalized. Dynamical interpolation between exact dynamics and rigidly constrained dynamics is demonstrated, with the associated gain in the time step stability. On the other hand, exact sampling with possible gain in the mixing time for Metropolized sampling methods is analyzed.

Furthermore, for the  $N$ -alkane model with large  $N$ , torsion angles are penalized, a large mass-penalty of order  $\mathcal{O}(N)$ , i.e.,  $\nu = \bar{\nu}N$ , where  $N$  is the system size is considered. For dynamics, it induces a gain in the time step stability region that grows with  $N$ , while numerical evidence is given that some macroscopic timescales, e.g., low frequencies of the chain length dynamics, remain of order  $\mathcal{O}(1)$ . For sampling, it induces a similar gain for mixing time in terms of iteration steps, and measured with autocorrelation of the chain length evolution. Rigorous proofs with explicit scalings of this behavior are provided for the case of a linear atomic chain with quadratic (harmonic) interactions in Section 7, and consistence of the IMMP macroscopic dynamics towards a stochastic wave equation when the re-scaled penalty vanishes  $\bar{\nu} \rightarrow 0$  is demonstrated.

**Acknowledgments:** The research of M.R. was partially supported by the EPSRC grant GR/S70883/01 while he was visiting Mathematics Institute, University of Warwick. The research of P.P. was partially supported by the National Science Foundation under the grant NSF-DMS-0813893 and by the Office of Advanced Scientific Computing Research, U.S. Department of Energy; the work was partly done at the ORNL, which is managed by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725.

**2. Langevin processes and sampling of canonical distribution.** We consider a Hamiltonian system in the phase-space  $\mathbb{R}^d \times \mathbb{R}^d$  with the Hamiltonian  $H$  in the form

$$H(p, q) = \frac{1}{2} p^T M^{-1} p + V(q), \quad (2.1)$$

We use generic matrix notation, for instance, the Euclidean scalar product of two vectors  $p_1, p_2 \in \mathbb{R}^N$  is denoted by  $p_1^T p_2$ , and the gradients of mappings from  $\mathbb{R}^d$  to  $\mathbb{R}^n$  with respect to standard bases are represented by matrices

$$(\nabla_q^T \xi)_{ij} = (\nabla_q \xi)_{ji} = \frac{\partial \xi_i}{\partial q_j}, \quad i = 1, \dots, n, \quad j = 1, \dots, d.$$

When the system is thermostatted, i.e., kept at the constant temperature, the long time distribution of the system in the phase-space is given by the canonical equilibrium measure at the inverse temperature  $\beta$  (also called the NVT distribution) given by

$$\mu(dp dq) = \frac{1}{Z} e^{-\beta H(p, q)} dp dq, \quad Z = \int_{\mathbb{R}^d \times \mathbb{R}^d} e^{-\beta H(p, q)} dp dq, \quad (2.2)$$

with the normalization constant  $Z < \infty$ . The standard dynamics used to model thermostatted systems are given by Langevin processes.

**DEFINITION 2.1** (Langevin process). *A Langevin process at the inverse temperature  $\beta$  with the Hamiltonian  $H(q, p)$ ,  $(p, q) \in \mathbb{R}^d \times \mathbb{R}^d$ , the  $d \times d$  dissipation matrix  $\gamma$ , and the fluctuation matrix  $\sigma$  is given by the stochastic differential equations*

$$\begin{cases} \dot{q} = \nabla_p H \\ \dot{p} = -\nabla_q H - \gamma \dot{q} + \sigma \dot{W}, \end{cases} \quad (2.3)$$

where  $\dot{W}$  is a standard white noise (Wiener process), and  $\sigma \in \mathbb{R}^d \times \mathbb{R}^d$  satisfies the fluctuation-dissipation identity

$$\sigma \sigma^T = \frac{2}{\beta} \gamma.$$

For any  $\gamma$ , the process is reversible with respect to the stationary canonical distribution (2.2). Furthermore, if  $\gamma$  is strictly positive definite, the process is ergodic.

Throughout the paper, stochastic integrands have finite variation thus the stochastic integration (e.g., Itô or Stratonovitch) need not be specified. Furthermore, the usual global Lipschitz conditions (see [32]) on  $H$  and  $\xi$  are assumed, ensuring well-posedness of the considered stochastic differential equations. The analysis presented in the paper can be generalized to a position dependent dissipation matrix  $\gamma = \gamma(q)$ .

The mapping  $\xi : \mathbb{R}^d \rightarrow \mathbb{R}^n$ , defines  $n \leq d$  degrees of freedom, given by smooth functions taking values in a neighborhood of 0. We assume that the mapping  $\xi$  is regular (i.e., with a non-degenerate Jacobian) in an open  $\delta$ -neighborhood  $\mathcal{O}_\delta = \{q \mid \|\xi(q)\| < \delta\}$  of  $\xi^{-1}(0)$ , hence defining a smooth sub-manifold of  $\mathbb{R}^d$  denoted  $\mathcal{M}_z = \xi^{-1}(z)$  for  $z$  in a neighborhood of the origin. The dependence of the potential  $V$  with respect to the degrees of freedom  $\xi$  is expected to be “stiff” in the second variable. In Section 6 we will introduce the stiffness parameter  $\epsilon$ . In that section we shall assume that such parameter dependence can be explicitly identified, and that the potential energy  $V$  can be written in the form

$$V(q) = U(q, \frac{\xi(q)}{\epsilon}), \quad (2.4)$$

where the function  $U : \mathbb{R}^d \times \mathbb{R}^n \rightarrow \mathbb{R}$  satisfies the coercivity condition  $\lim_{z \rightarrow +\infty} U(q, z) = +\infty$ . The fast degrees of freedom  $\xi$  of states at a given energy then remain in a closed neighborhood of the origin as the stiffness parameter  $\epsilon \rightarrow 0$ . In this limit the system is confined to the sub-manifold  $\mathcal{M}_0$  which is usually called the “slow manifold”.

**3. The implicit mass-matrix penalization method.** In this section we focus on properties of the IMMP method. The multiscale structure of the potential  $V$  need not be known in order to apply the method. Thus, in this section, we consider the potential  $V$  in the form where we do not impose the structural assumption (2.4) on the potential function  $V : \mathbb{R}^d \rightarrow \mathbb{R}$ .

**3.1. Description of the method.** The new, penalized mass-matrix of the system is the position dependent tensor defined in (1.3). The associated modified impulses are denoted

$$p_\nu = M_\nu(q) M^{-1} p. \quad (3.1)$$

When  $\nu$  becomes large, the velocities are bound to remain tangent to the manifolds  $\{q \mid \xi(q) = z\}$ , and orthogonal motions are arbitrarily slowed down. Conversely, when  $\nu = 0$ , one recovers the original highly oscillatory system. Since the modification in  $M_\nu$  depends on the position  $q$ , new geometry is introduced and an additional correction (1.4) in the potential energy is required in order to preserve original statistics in the position variable. This correction is in fact close to the standard Fixman corrector for  $\nu$  large (see (6.5)). Defining  $G(q)$  as the  $n \times n$  Gram matrix associated with the fast degrees of freedom

$$G(q) = \nabla_q^T \xi M^{-1} \nabla_q \xi, \quad (3.2)$$

one has the following property of the correcting potential.

**PROPOSITION 3.1.** *Up to an additive constant, we have*

$$V_{\text{fix},\nu}(q) = \frac{1}{2\beta} \ln \det \left( G(q) + \frac{1}{\nu^2} M_z^{-1} \right), \quad (3.3)$$

and thus (up to additive constants)

$$\lim_{\nu \rightarrow +\infty} V_{\text{fix},\nu} = V_{\text{fix}} = \frac{1}{2\beta} \ln \det (G(q)), \quad \text{and} \quad \lim_{\nu \rightarrow 0} V_{\text{fix},\nu} = 0.$$

*Proof.* Using the identity for a non-diagonal matrix  $J$  of dimension  $n_1 \times n_2$ :

$$\det (\text{Id}_{n_1} + J J^T) = \det (\text{Id}_{n_2} + J^T J),$$

one observes

$$\det(M_\nu) = \det(M) \det(\nu^2 M_z) \det(G + \frac{1}{\nu^2} M_z^{-1})$$

from which the expression for the corrected Fixman potential follows.  $\square$

The associated modified Hamiltonian is then given by

$$H_\nu(p_\nu, q) = \frac{1}{2} p_\nu^T M_\nu^{-1} p_\nu + V(q) + V_{\text{fix},\nu}(q), \quad (3.4)$$

and  $H_0 = H$  is the original Hamiltonian (1.1).

Sampling such a system can be done using the standard Langevin stochastic perturbation as detailed in Definition 2.1. However, the direct discretization of the equation of motion given by  $H_\nu$  (e.g., by an explicit scheme) is bound to be unstable from non-linear instabilities when the fast degrees of freedom are not affine functions. In order to construct stable schemes one may rather use an implicit formulation of the Hamiltonian (3.4), in conjunction with a solver which enforces the constraints. To obtain such a formulation we extend the state space with  $n$  new variables  $(z_1, \dots, z_n)$ , and associated momenta  $(p_{z_1}, \dots, p_{z_n})$ . The auxiliary mass-matrix for the new degrees of freedom is then given by  $M_z$ . The new extended Hamiltonian of the system  $H_{\text{IMMP}}$ , defined by (1.5), is now defined in  $\mathbb{R}^{d+n} \times \mathbb{R}^{d+n}$ , where  $n$  position constraints denoted by  $(C_\nu)$  are included. This construction implies  $n$  hidden constraints on momenta. The equivalence of the two Hamiltonians (3.4) and (1.5) formulations is stated as a simple separate lemma.

LEMMA 3.2. *The equations of motion associated with the penalized mass-matrix Hamiltonian (3.4) or the extended Hamiltonian with constraints (1.5) are identical.*

*Proof.* The Lagrangian associated with  $H_{\text{IMMP}}$  is given by

$$L_{\text{IMMP}}(\dot{q}, \dot{z}, q, z) = \frac{1}{2} \dot{q}^T M \dot{q} + \frac{1}{2} \dot{z}^T M_z \dot{z} - V(q) - V_{\text{fix},\nu}(q),$$

and includes hidden constraints on velocities  $\dot{z} = \nu \nabla_q^T \xi \dot{q}$  implied by the constraints  $(C_\nu)$  on position variables. Replacing  $\dot{z}$  and  $z$  in  $L_{\text{IMMP}}$  by their expressions as functions of  $\dot{q}$  and  $q$ , one obtains the Lagrangian associated with  $H_\nu$ .  $\square$

The stochastically perturbed equations of motion of the Langevin type associated with (1.5) define the dynamics with implicit mass-matrix penalization.

DEFINITION 3.3 (IMMP). *The implicit Langevin process associated with Hamiltonian  $H_{\text{IMMP}}$  and constraints  $(C_\nu)$  is defined by the following equations of motion*

$$\begin{cases} \dot{q} = M^{-1} p \\ \dot{z} = M_z^{-1} p_z \\ \dot{p} = -\nabla_q V(q) - \nabla_q V_{\text{fix},\nu}(q) - \gamma \dot{q} + \sigma \dot{W} - \nabla_q \xi \dot{\lambda} \\ \dot{p}_z = -\gamma_z \dot{z} + \sigma_z \dot{W}_z + \frac{\dot{\lambda}}{\nu} \\ \xi(q) = \frac{z}{\nu}, \end{cases} \quad (3.5) \quad (C_\nu)$$

*The process  $\dot{W}$  (resp.  $\dot{W}_z$ ) is a standard multi-dimensional white noise,  $\gamma$  (resp.  $\gamma_z$ ) a  $d \times d$  (resp.  $n \times n$ ) non-negative symmetric dissipation matrix,  $\sigma$  (resp.  $\sigma_z$ ) is the fluctuation matrix satisfying  $\sigma \sigma^T = \frac{2}{\beta} \gamma$  (resp.  $\sigma_z \sigma_z^T = \frac{2}{\beta} \gamma_z$ ). The processes  $\lambda \in \mathbb{R}^n$  are Lagrange multipliers associated with the constraints  $(C_\nu)$  and adapted with the white noise.*

This process is naturally equivalent to the explicit mass-penalized Langevin process in  $\mathbb{R}^d \times \mathbb{R}^d$  associated with  $M_\nu$ . Moreover, when the penalization vanishes ( $\nu \rightarrow 0$ ), the evolution law of the process  $\{p_t, q_t\}_{t \geq 0}$  or  $\{(p_\nu)_t, q_t\}_{t \geq 0}$  converges towards the original dynamics.

**PROPOSITION 3.4.** *The stochastic process with constraints (3.5) is well-posed and equivalent to the Langevin diffusion in  $\mathbb{R}^d \times \mathbb{R}^d$  (see Definition 2.1), with the mass-penalized Hamiltonian  $H_\nu$  (3.4), and the dissipation matrix given by*

$$\gamma_\nu(q) = \gamma + \nu^2 \nabla_q \xi \gamma_z \nabla_q^T \xi.$$

*Furthermore, the process is reversible and ergodic with respect to the canonical distribution (3.8) (with marginal in position variables given by the original potential, i.e., up to the normalization,  $e^{-\beta V(q)} dq$ ).*

*Proof.* Imposing the constraints implies  $\nabla_q^T \xi M^{-1} p = \frac{1}{\nu} M_z^{-1} p_z$ . Thus by the definition of  $p_\nu$  we have

$$p_\nu = p + \nu \nabla_q \xi p_z.$$

Since the position process  $\{q_t\}_{t \geq 0}$  is of finite variation, a short computation shows that for each coordinate  $i = 1, \dots, d$

$$\dot{p}_\nu^i = \dot{p}^i + \nu \partial_{q_i} \xi \dot{p}_z + \nu^2 \dot{q}^T \nabla_q (\partial_{q_i} \xi) p_z. \quad (3.6)$$

Furthermore,

$$-\partial_{q_i} \left( \frac{1}{2} p_\nu^T M_\nu^{-1} p_\nu \right) = \partial_{q_i} \left( \frac{1}{2} \dot{q} M_\nu \dot{q} \right) = \nu^2 \dot{q}^T \nabla_q (\partial_{q_i} \xi) M_z \nabla_q^T \xi \dot{q},$$

and thus

$$\dot{p}_\nu = \dot{p} + \nu \nabla_q \xi \dot{p}_z - \nabla_q \left( \frac{1}{2} p_\nu^T M_\nu^{-1} p_\nu \right).$$

Substituting the expressions for  $\dot{p}$  and  $\dot{p}_z$  from (3.5) into (3.6) we obtain

$$\dot{p}_\nu = -\frac{1}{2} p_\nu^T \nabla_q M_\nu^{-1} p_\nu - \nabla_q V(q) - \nabla_q V_{\text{fix},\nu}(q) - \gamma \dot{q} - \nu \nabla_q \xi \gamma_z \dot{z} + \sigma \dot{W} + \nu \nabla_q \xi \sigma_z \dot{W}_z, \quad (3.7)$$

which yields the result.  $\square$

**3.2. Exact sampling in position variables.** By construction, statistics of positions  $q$  of the mass penalized Hamiltonian are independent of the penalization, leading to the *exact canonical statistics* in position variables.

**PROPOSITION 3.5** (Exact statistics). *The canonical distribution associated with the mass-penalized Hamiltonian (3.4) is given by*

$$\mu_\nu(dp_\nu dq) = \frac{1}{Z_\nu} e^{-\beta H_\nu(p_\nu, q)} dp_\nu dq. \quad (3.8)$$

*Its marginal probability distribution in  $q$  is*

$$\frac{1}{Z_\nu} \int e^{-\beta H_\nu(p_\nu, q)} dp_\nu = \frac{e^{-\beta V(q)} dq}{\int e^{-\beta V(q)} dq}$$

*which is the original canonical distribution (2.2) in the position variables, and is independent of the mass penalization parameter  $\nu$ .*

*Proof.* The normalization of Gaussian integrals in the  $p_\nu$  variables yields

$$\int e^{-\beta \frac{1}{2} p_\nu^T M_\nu^{-1} p_\nu} dp_\nu = \left( \frac{2\pi}{\beta} \right)^{d/2} \sqrt{\det(M_\nu)},$$

which is cancelled out by the Fixman corrector  $V_{\text{fix},\nu}$  and the result follows.  $\square$

**3.3. Interpolation between exact and constrained dynamics.** In this section, the IMMP dynamics is shown to be an interpolation between exact dynamics ( $\nu = 0$ ), and constrained dynamics ( $\nu = +\infty$ ).

**PROPOSITION 3.6** (Small penalty). *When  $\nu \rightarrow 0$  the evolution law of the processes  $\{p_t, q_t\}_{t \geq 0}$  or  $\{(p_\nu)_t, q_t\}_{t \geq 0}$  defined by the implicit equations (3.5) converges (in the sense of probability distributions on continuous paths endowed with the uniform convergence) towards the process solving the original Langevin dynamics (2.3).*

*Proof.* The stochastic differential equation defined by  $\dot{q} = M_\nu^{-1} p_\nu$  and (3.7) has smooth coefficients which depend on  $\nu$  in a continuous fashion ( $\nu \mapsto M_\nu$  and  $\nu \mapsto V_{\text{fix},\nu}$  are continuous). Standard results on weak convergence ([16]) of stochastic processes imply the result as stated.  $\square$

When the mass penalty tends to infinity, the IMMP process converges to a constrained process on the manifold  $\mathcal{M}_{z_{t=0}} = \{q \mid \xi(q) = z_{t=0}\}$ .

**PROPOSITION 3.7** (Large penalty). *Consider a family of initial conditions indexed by  $\nu$  and satisfying*

$$\sup_\nu |\nu (\xi(q_{t=0}) - z_{t=0})| < +\infty,$$

*and assume that the Gram matrix  $G$  is invertible in a neighborhood of  $\mathcal{M}_{z_{t=0}}$ . Then when  $\nu \rightarrow +\infty$  the IMMP Langevin stochastic process (3.5) converges in distribution towards the decoupled limiting processes with constraints*

$$\begin{cases} \dot{q} = M^{-1} p, \\ \dot{p} = -\nabla_q V - \nabla_q V_{\text{fix}} - \gamma \dot{q} + \sigma \dot{W} - \nabla_q \xi \dot{\lambda}, \\ \xi(q) = z_{t=0}, \\ \dot{z} = M_z^{-1} p_z, \\ \dot{p}_z = -\gamma_z \dot{z} + \sigma_z \dot{W}_z. \end{cases} \quad (C) \quad (3.9)$$

where  $\{\lambda_t\}_{t \geq 0}$  are adapted stochastic processes defining the Lagrange multipliers associated with the constraints (C).

Furthermore, the process  $\{q_t, p_t\}_{t \geq 0}$  defines an effective dynamics with constraints (see also Definition 6.2) on the sub-manifold  $\mathcal{M}_{z_{t=0}}$ . It is reversible with respect to its stationary canonical distribution given, up to the normalization, by the “stiff” Boltzmann distribution

$$e^{-\beta(H(p,q) + V_{\text{fix}}(q))} \sigma_{T^* \mathcal{M}_{z_{t=0}}} (dp dq)$$

with the  $q$ -marginal  $e^{-\beta V(q)} \delta_{\xi(q)=0}(dq)$ . When  $\gamma$  and  $\gamma_z$  are strictly positive definite the process is ergodic.

*Proof.* By a simple translation, it is sufficient to show the proposition for  $z_{t=0} = 0$ . Satisfying the constraint (C $_\nu$ ) in (3.5) implies a hidden constraint in the momentum space,  $\nabla_q \xi M^{-1} p = \frac{1}{\nu} M_z^{-1} p_z$ . Differentiating this expression with respect to time and replacing the result in (3.5) yields an explicit formula for the Lagrange multipliers

$$\dot{\lambda} = (G + \frac{1}{\nu^2} M_z^{-1})^{-1} \left[ \text{Hess}(\xi) (M^{-1} p, M^{-1} p) + \nabla_q \xi M^{-1} f_q - \frac{1}{\nu} M_z^{-1} f_z \right], \quad (3.10)$$

with forces  $(f_q, f_z)$

$$\begin{aligned} f_q &= -\nabla_q V - \nabla_q V_{\text{fix},\nu} - \gamma M^{-1} p + \sigma \dot{W}, \\ f_z &= -\gamma_z M_z^{-1} p_z + \sigma_z \dot{W}_z, \end{aligned}$$

and the Hessian  $\text{Hess}(\xi)$  of the mapping  $\xi$  acting on the velocities  $M^{-1} p$ . This calculation shows that (3.5) is in fact a standard stochastic differential equation with smooth coefficients, and thus has a unique strong

solution. The coefficients of these stochastic differential equations are continuous in the limit  $\frac{1}{\nu} \rightarrow 0$ , at least in a  $\delta$ -neighborhood of  $\mathcal{M}_0$  in which  $G$  is invertible. The formally computed limiting process is given by (3.9) with the Lagrange multipliers solving

$$\dot{\lambda} = G^{-1} (\text{Hess}(\xi) (M^{-1}p, M^{-1}p) + \nabla_q^T \xi M^{-1} f_q) .$$

By construction, this limiting process satisfies the constraint  $\xi(q) = 0$ . Its coefficients are Lipschitz and the process is well-posed. As a result of those properties, the rigorous proof of weak convergence follows classical arguments, see [16], that are divided into three steps

- (i) We truncate the process (3.5) to a compact neighborhood of  $\mathcal{M}_0$ .
- (ii) The continuity of the Markov generator with respect to  $\frac{1}{\nu}$  implies tightness for the associated  $\frac{1}{\nu}$ -sequences of truncated processes with the limit being uniquely defined by (3.9).
- (iii) The limiting process remains on  $\mathcal{M}_0$ , which implies weak convergence of the sequence without truncation.

The process (3.9) is thus a Langevin process with constraints, exhibiting reversibility properties with respect to the associated Boltzmann canonical measure and is ergodic when  $\gamma$  is strictly positive definite (see the summary in Appendix B). Note that the  $q$ -marginal is geometrically corrected by the Fixman potential term.  $\square$

We conclude this section by discussing some consequences for numerical computations.

**REMARK 3.8.** Proposition 3.6 and 3.7 imply that the IMMP scheme is a tunable interpolation between the exact stochastic dynamics (2.3), and the stochastic dynamics with constraints (3.9). If one prefers to interpolate with rigidly constrained dynamics (i.e., without the Fixman correction, see also Section 6 for a detailed discussion on “stiff”, as opposed to “rigid”, constrained dynamics), one removes the Fixman correction in the force evaluation.

**4. Numerical integration.** The key ingredient for achieving efficient numerical simulation is to use an integrator that enforces the constraints associated with the implicit formulation of the mass penalized dynamics (3.5). The implicit structure of (3.5) leads to numerical schemes that are potentially asymptotically stable in stiff cases (Section 6). On the other hand, when the penalization  $\nu$  vanishes with the time-step the scheme becomes consistent with respect to the original exact dynamics (2.3). One may then consider the mass-penalization introduced here as a special method of pre-conditioning for a stiff ODE system with an “implicit”, in the time evolution sense, structure. Here, the “implicit” structure amounts to solving the imposed constraints  $\xi(q) = z/\nu$  in (3.5).

It lies outside the scope of this paper to review standard numerical methods for constrained mechanical systems, we refer to [14] as a classical textbook, and to the series ([40, 25, 15, 21, 8]) as a sample of works on practical developments of numerical methods. The IMMP method is presented with the classical leapfrog/Verlet scheme that enforces constraints, usually called RATTLE in (4.1). It can be implemented by a simple modification of standard schemes constraining fDOFs. The scheme is second order, reversible and symplectic. This choice is largely a presentation matter, for practical purposes one can refer to one’s favorite numerical integrator for Hamiltonian systems with or without stochastic perturbations.

For accurate sampling of the equilibrium distribution, one can also add a Metropolis acceptance, rejection time-step corrector at each time step of the deterministic integrator. If the underlying integrator is reversible and preserves the phase-space measure, this extension leads to a scheme which *exactly* preserves canonical distributions. The Metropolis correction is used in Hybrid Monte-Carlo (HMC) methods, which are sampling algorithms relying on the underlying dynamics of the system to generate moves in the configuration space, which are accepted or rejected according to the Metropolis rule. The Metropolis acceptance/rejection step can be used at each integration time step of the Langevin process, usually referred to as Generalized Hybrid Monte Carlo (GHMC) introduced in [22]. However, the necessity of the momentum flip when a rejection occurs destroys the dynamical features of the Langevin process when the rate of rejection does not vanish, and makes the latter rather behave similarly to an overdamped dynamics. Note also that for usual schemes integrating the Hamiltonian dynamics, the average acceptance ratio

often decreases when the dimension of the system increases ([24]). Many improvements and modifications ([11, 22, 2, 23]) of the HMC algorithm have been developed since its introduction in [12] for simulations applied to quantum statistical field theories. Subsequently it has been also employed to a wide range of simulations in macromolecular systems, e.g., [9, 38].

We implement numerical discretization of the Langevin process with constraints (3.5) obtained by splitting the Hamiltonian part and the Gaussian fluctuation/dissipation perturbation. Note that by using a Metropolis acceptance/rejection rule (HMC), or simply by weighting statistical averages, the forces associated with the Fixman corrector need not be computed when one is interested in sampling only.

**4.1. The IMMP integrator.** We recall that we consider the IMMP dynamics (3.5), which consists of the following elements

1. the Hamiltonian  $H_{\text{IMMP}}$  defined in (1.5), which defines the deterministic dynamics of the IMMP.
2. the dissipation matrix  $\text{diag}(\gamma, \gamma_z)$ , and the inverse temperature  $\beta$ , which defines the features of the stochastic thermostat.

SCHEME 4.1 (Dynamical integrator).

*Step 1: Integrate the Hamiltonian part with:*

$$\begin{aligned}
 & \begin{cases} p_{n+1/2} = p_n - \frac{\delta t}{2}(\nabla_q V + \nabla_q V_{\text{fix},\nu})(q_n) - \nabla_q \xi(q_n) \lambda_{n+1/2} \\ p_{n+1/2}^z = p_n^z + \frac{1}{\nu} \lambda_{n+1/2} \end{cases} \\
 & \begin{cases} q_{n+1} = q_n + \delta t M^{-1} p_{n+1/2} \\ z_{n+1} = z_n + \delta t M_z^{-1} p_{n+1/2}^z \end{cases} \\
 & \xi(q_{n+1}) = \frac{z_{n+1}}{\nu} \tag{C_{1/2}}
 \end{aligned} \tag{4.1}$$

$$\begin{aligned}
 & \begin{cases} p_{n+1} = p_{n+1/2} - \frac{\delta t}{2}(\nabla_q V + \nabla_q V_{\text{fix},\nu})(q_{n+1}) - \nabla_q \xi(q_{n+1}) \lambda_{n+1} \\ p_{n+1}^z = p_{n+1}^z + \frac{1}{\nu} \lambda_{n+1} \end{cases} \\
 & \nabla_q^T \xi(q_{n+1}) M^{-1} p_{n+1} = \frac{1}{\nu} M_z^{-1} p_{n+1}^z \tag{C_1}.
 \end{aligned}$$

*Step 2: Integrate if necessary the Gaussian fluctuation/dissipation part with a mid-point Euler scheme with constraints (see Appendix C).*

Here again, one can remove the Fixman correction forces  $\nabla_q V_{\text{fix},\nu}$  in force evaluation in order to interpolate with usual rigid constraints dynamics.

Note a useful variant of the integrator, which occurs when the potential dependence with respect to the penalized variables is known explicitly  $V(q) = U(q, \xi(q))$ . In such a case, the expressions (4.1) is replaced

by

$$\begin{cases} p_{n+1/2} = p_n - \frac{\delta t}{2}(\nabla_1 U + \nabla_q V_{\text{fix},\nu})(q_n) - \nabla_q \xi(q_n) \lambda_{n+1/2} \\ p_{n+1/2}^z = p_n^z - \frac{\delta t}{2\nu} \nabla_2 U(q_n) + \frac{1}{\nu} \lambda_{n+1/2} \\ q_{n+1} = q_n + \delta t M^{-1} p_{n+1/2} \\ z_{n+1} = z_n + \delta t M_z^{-1} p_{n+1/2}^z \\ \xi(q_{n+1}) = \frac{z_{n+1}}{\nu} \end{cases} \quad (C_{1/2})$$

$$\begin{cases} p_{n+1} = p_{n+1/2} - \frac{\delta t}{2}(\nabla_1 U + \nabla_q V_{\text{fix},\nu})(q_{n+1}) - \nabla_q \xi(q_{n+1}) \lambda_{n+1} \\ p_{n+1}^z = p_{n+1/2}^z - \frac{\delta t}{2\nu} \nabla_2 U(q_{n+1}) + \frac{1}{\nu} \lambda_{n+1} \\ \nabla_q^T \xi(q_{n+1}) M^{-1} p_{n+1} = \frac{1}{\nu} M_z^{-1} p_{n+1}^z \end{cases} \quad (C_1).$$

where in the above,  $\nabla_1$  and  $\nabla_2$  denote the derivatives with respect to the first and the second variables, respectively. This variant consists in applying a part of the force to the auxiliary variables instead of directly to the system. For penalized degrees of freedom that move far from their equilibrium value, e.g., torsion angles, see Section 5, this may increase the stability of the algorithm enforcing the constraints. In the same way for slow/fast systems (see Section 6.2), the scheme (4.2) will lead to asymptotic stability in the large stiffness limit.

**4.2. Exact sampling and Monte-Carlo scheme.** To obtain exact sampling by correcting time step errors and, if necessary, the geometric bias, a Metropolis acceptance/rejection is added. A domain of phase-space  $D_{\delta t} \subset \mathbb{R}^{2d} \times \mathbb{R}^{2n}$  where the integrator (4.1) with constraints has a unique solution is also considered. In practice this set is simply the set of configurations for which the algorithm used to enforce constraints in (4.1) (e.g., a Newton algorithm) converges in a given number of steps. In practice, the (equilibrium) probability for the system of lying outside  $D_{\delta t}$  goes to zero exponentially fast with the time step, and the latter is taken such that this probability is negligible.

SCHEME 4.2 (Leapfrog/Verlet algorithm with Metropolis correction for Langevin IMMP (3.5)).

*Step 1: Compute  $(q_{n+1}, z_{n+1}, p_{n+1}, p_{n+1}^z)$  with the integrator (4.1), and set*

$$\Delta H_{n+1} = H_{\text{IMMP}}(q_{n+1}, z_{n+1}, p_{n+1}, p_{n+1}^z) - H_{\text{IMMP}}(q_n, z_n, p_n, p_n^z).$$

*If  $(q_{n+1}, z_{n+1}, p_{n+1}, p_{n+1}^z)$  does not belong to  $D_{\delta t}$ , set  $\Delta H_{n+1} = +\infty$ .*

*Step 2: Accept the step with the probability  $\min(1, e^{-\beta \Delta H_{n+1}})$ , otherwise reject, flip momenta, and set*

$$(q_{n+1}, z_{n+1}, p_{n+1}, p_{n+1}^z) = (q_n, z_n, -p_n, -p_n^z).$$

*Step 3: Integrate the Gaussian fluctuation/dissipation part with a mid-point Euler scheme (for details see Appendix C).*

REMARK 4.1. A useful variant, when using HMC strategies, consists in modifying the Hamiltonian  $H_{\text{IMMP}}$  (1.5) in the integrator (importance sampling) by neglecting the Fixman corrector

$$\begin{cases} \tilde{H}_{\text{IMMP}}(p, p_z, q, z) = \frac{1}{2} p^T M^{-1} p + \frac{1}{2} p_z^T M_z^{-1} p_z + \tilde{V}(q) \\ \xi(q) = \frac{z}{\nu} \end{cases} \quad (C_\nu),$$

where  $\tilde{V}$  is a potential that may be chosen arbitrarily (typically  $\tilde{V} = V$ ). Indeed, only the underlying phase space structure, which does not depend on  $\tilde{V}$ , is necessary in HMC methods. The correct potential  $V + V_{\text{fix},\nu}$

has to be used in the Metropolis step only, or alternatively by weighting ensemble averages, in order to ensure exact canonical sampling. Thus potentially costly evaluations of the gradient of the Fixman corrector  $V_{\text{fix},\nu}$  can be avoided. This numerically constructed Markov chain preserves the canonical distribution.

**PROPOSITION 4.2** (Exact sampling). *The Monte-Carlo algorithm generated from (3.5) as described in Scheme 4.2 generates a Markov chain that leaves invariant the canonical distribution (3.8) (conditioned in the constraints stability domain  $D_{\delta t}$  of (4.1)). The marginal in position variables of the distribution (3.8) is the original distribution  $e^{-\beta V(q)}dq$  which is independent of the mass-penalization  $\nu$ .*

*Proof.* The statement follows from reversibility and measure preserving properties of Verlet schemes (see [14]), from the Hybrid Monte Carlo rule ([13, 24]), and from the construction of the mass-penalized Hamiltonian (Proposition 3.4).  $\square$

**4.3. Interpolation between exact and constrained dynamics.** First we study the limit towards the leapfrog/Verlet scheme.

**PROPOSITION 4.3** (Small penalty limit). *Assume that the IMMP integrator (4.1) or (4.2) is locally well-defined, and that  $\nu \rightarrow 0$ . Then the latter two (computed with or without the Fixman correcting forces  $\nabla_q V_{\text{fix},\nu}$ ) converge at order  $\mathcal{O}(\nu^2)$  towards the leapfrog/Verlet scheme for the exact dynamics.*

*Proof.* Consider the auxiliary variables in the scaling  $(\bar{z}, \bar{p}^z) = (\frac{z}{\nu}, \frac{p^z}{\nu})$ , and the following shift of Lagrange multipliers in (4.2), or equivalently in (4.1)

$$\lambda \rightarrow -\nu^2 \lambda + \frac{\delta t}{2} \nabla_2 U,$$

then the scheme (4.2) becomes

$$\begin{cases} p_{n+1/2} = p_n - \frac{\delta t}{2} (\nabla_q U + \nabla_q V_{\text{fix},\nu})(q_n) + \nu^2 \nabla_q \xi(q_n) \lambda_{n+1/2} \\ \bar{p}_{n+1/2}^z = \bar{p}_n^z - \lambda_{n+1/2} \\ \begin{cases} q_{n+1} = q_n + \delta t M^{-1} p_{n+1/2} \\ \bar{z}_{n+1} = \bar{z}_n + \delta t M_z^{-1} \bar{p}_{n+1/2}^z \end{cases} \\ \xi(q_{n+1}) = \bar{z}_{n+1} \end{cases} \quad (C_{1/2})$$

$$\begin{cases} p_{n+1} = p_{n+1/2} - \frac{\delta t}{2} (\nabla_1 U + \nabla_q V_{\text{fix},\nu})(q_{n+1}) + \nu^2 \nabla_q \xi(q_{n+1}) \lambda_{n+1} \\ \bar{p}_{n+1}^z = \bar{p}_{n+1/2}^z - \lambda_{n+1} \\ \nabla_q^T \xi(q_{n+1}) M^{-1} p_{n+1} = M_z^{-1} \bar{p}_{n+1}^z \end{cases} \quad (C_1),$$
(4.4)

which converges at order  $\mathcal{O}(\nu^2)$  to a decoupled scheme where the  $(q, p)$  variables evolve according to the usual leapfrog/Verlet scheme, and the auxiliary variables  $(\bar{z}, \bar{p}^z)$  are enforced by the constraints  $\xi(q) = \bar{z}$ . Now from (1.3),  $V_{\text{fix},\nu}$  is of order  $\mathcal{O}(\nu^2)$  which completes the proof.  $\square$

One can now construct schemes consistent with respect to the exact dynamics by letting the penalty  $\nu = \bar{\nu} \delta t^k$  go to zero with the time step, for some  $k > 0$ . Indeed, Proposition 3.6 shows that the mass-penalized dynamics (3.5) converges towards the exact original dynamics for  $\nu = 0$  at order  $\mathcal{O}(\nu^2)$ . Consequently most of the usual numerical schemes will be consistent at their own approximation order but bounded above by  $2k$ . The order of convergence refers to the maximal integer  $k$  such that the convergence of trajectories with respect to the uniform norm occurs at the rate  $\mathcal{O}(\delta t^k)$ . Neglecting the order of the fluctuation/dissipation part (Step 2 of Scheme 4.1), we deduce the following consistency property.

**PROPOSITION 4.4** (Time-step consistency with exact dynamics). *Assume that the integrator (4.1) or (4.2) are locally well-defined, and that  $\nu = \bar{\nu} \delta t$ . Then the IMMP numerical scheme (4.1) or (4.2) (computed with or without the Fixman correcting forces  $\nabla_q V_{\text{fix},\nu}$ ), is of the order 2 and consistent with respect to the original exact deterministic dynamics (2.3) (i.e., with  $\gamma = 0$ ,  $\sigma = 0$ ).*

*Proof.* Following the proof of Proposition 4.3 we have, by the implicit function theorem, locally, the RATTLE scheme is a standard leapfrog scheme (see [14, 36]), and the associated local mapping depends continuously on  $\nu^2$  when  $\nu \rightarrow 0$ . Therefore the usual calculation of the order of the leapfrog scheme (see [14]) holds uniformly with respect to  $\nu$ , and (4.1) or (4.2) are of order 2 consistent uniformly in  $\nu$ . Now, the IMMP Hamiltonian (4.2) is a  $\nu^2$  perturbation of the exact Hamiltonian, and the result then follows from applying a simple Gronwall argument.  $\square$

Similarly, we also obtain the limit towards the constrained/RATTLE scheme.

**PROPOSITION 4.5** (Large penalty limit). *Assume that the integrators (4.1) or (4.2) are locally well-defined, and that  $\nu \rightarrow +\infty$  with  $\frac{z_0}{\nu} \rightarrow \bar{z}_0$ . Then the IMMP numerical scheme (4.1) or (4.2) (computed with or without the Fixman correcting forces  $\nabla_q V_{\text{fix},\nu}$ ) converges at order  $\frac{1}{\nu^2}$  towards the constrained/RATTLE scheme for the rigid constraints  $\xi(q_{n+1}) = \bar{z}_0$  (also computed with or without the associated geometric correcting Fixman forces  $\nabla_q V_{\text{fix}}$  defined by (3.3)).*

*Proof.* Following similar steps as in the proof of Proposition 4.3, with auxiliary variables in the scaling  $(\bar{z}, \bar{p}^z) = (\frac{z}{\nu}, \frac{p^z}{\nu})$ , and the following shift of Lagrange multipliers:

$$\lambda \rightarrow \lambda + \frac{\delta t}{2} \nabla_2 U,$$

we obtain convergence at the order  $\nu^{-2}$  to a decoupled scheme where the  $(q, p)$  variables evolve according to the usual constrained/RATTLE scheme. Now from (3.3),  $V_{\text{fix},\nu}$  is a perturbation of  $V_{\text{fix}}$  of order  $\nu^{-2}$  which completes the proof.  $\square$

**REMARK 4.6.** We conclude this section with a practical recipe for tuning the mass matrix penalty. Identifying a suitable value of  $\nu$  can be done, for instance, by computing the time averaged energy error (5.1) in a dynamical simulation, or the Metropolis acceptance ratio (5.2) in a sampling simulation, which gives a precise quantification of the time-step error. Then increasing the penalty  $\nu$  can save computational time as long as it leads to a reduction of the time-step error. Indeed, this means that the selected fDOFs are limiting the time-step stability region. Prescribing the time-step error, a maximal time-step  $\delta t_{\max}$  associated with the largest penalty  $\nu_{\max}$  that is able to improve stability can be obtained in this way. Finally, one can set, for example,  $\nu = \frac{\nu_{\max}}{\delta t_{\max}} \delta t$  in Scheme 4.1 to obtain an order two convergent scheme with an increased stability region.

**5. Numerical simulations for the  $N$ -alkane model.** The IMMP method is numerically tested on the united atom  $N$ -alkane model. The united atom model is a coarse-grained description of linear alkane isomers in which the hydrogen atoms are not resolved and the molecule is modeled by a chain of particles interacting with effective potentials, see, e.g., [30]. The integrator studied in the present section is described in Section 3 and Section 4, and is systematically compared with the exact dynamics, which is numerically integrated by a simple Leapfrog/Verlet scheme with or without a thermostat, and with the constrained dynamics numerically integrated by the RATTLE scheme with or without a thermostat. Constraints are resolved using a simple Newton algorithm with a Gaussian linear solver. The Fixman forces are not resolved in the dynamical part, and thus the dynamics interpolates for a large penalty with the “rigidly” constrained dynamics. However, they are used in the Metropolis rule when sampling is considered.

The Metropolis/HMC step for the dynamics integrator as described by Scheme 4.2 is added only when sampling is studied.

	Verlet	$\nu = 0.5$	$\nu = 1.0$	$\nu = 1.3$	$\nu = 1.9$	RATTLE
$\delta t_c^{\text{dyn}}$	.024 (1)	.032 (1)	.046 (1)	.059 (1)	.077 (1)	.093 (2)
$\delta t_c^{\text{samp}}$	.013 (1)	.014 (1)	.022 (1)	.028 (1)	.035 (1)	.049 (1)

TABLE 5.1

*Critical time steps  $t_c^{\text{dyn}}$  and  $\delta t_c^{\text{samp}}$  of the butane deterministic dynamics and of the sampling schemes, respectively. Note the interpolation property from constrained dynamics to exact dynamics.*

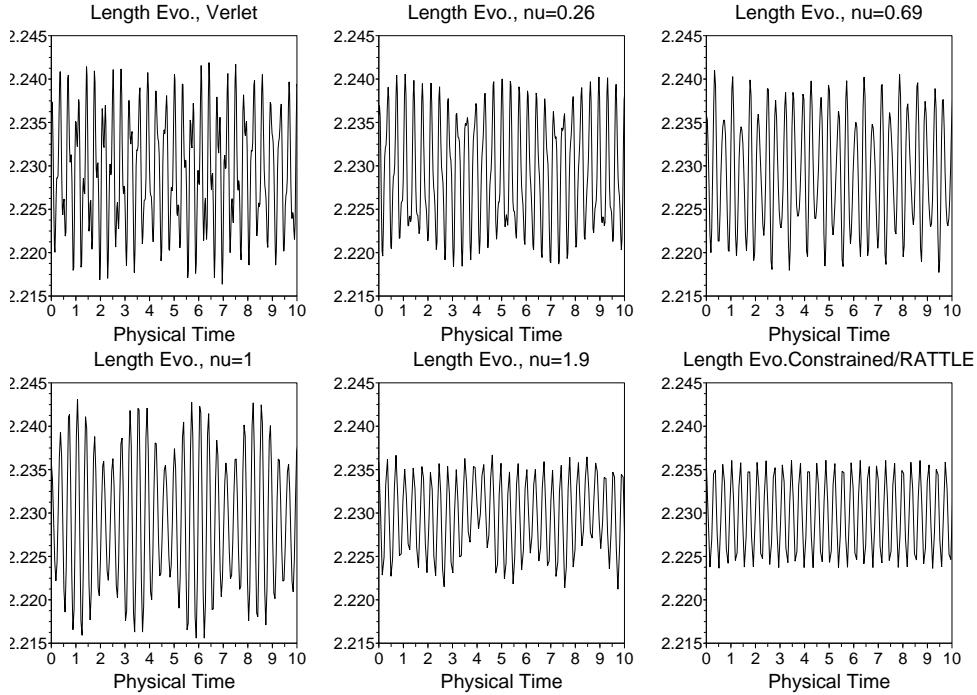


FIGURE 5.1. Oscillations of the butane end-to-end length, for the Verlet integrator, the IMMP integrator, and the constrained integrator. Simulations are computed with a prescribed initial energy. Note the interpolation property of the IMMP scheme. This figure is associated to the frequency analysis in Figure 5.2.

**5.1. The  $N$ -alkane model.** The model consists of a chain of  $N$  atoms with position vectors  $q^{(i)} \in \mathbb{R}^3$ , hence the vector of DOFs is  $q = (q^{(1)}, \dots, q^{(N)}) \in \mathbb{R}^{3N}$ . The mass of particles is normalized to be  $m^{(i)} = 1$  for all  $i = 1, \dots, N$ . The interaction potential  $V$  consists of three short-range potentials that involve 2-body, 3-body, and 4-body terms

$$V(q) = \sum_{i,j \in \mathcal{I}_{\text{bond}}} V_{\text{bond}}(q^{(i)}, q^{(j)}) + \sum_{i,j,k \in \mathcal{I}_{\text{angle}}} V_{\text{angle}}(q^{(i)}, q^{(j)}, q^{(k)}) + \sum_{i,j,k,l \in \mathcal{I}_{\text{torsion}}} V_{\text{torsion}}(q^{(i)}, q^{(j)}, q^{(k)}, q^{(l)}).$$

In the presented simulations we focus on short-range interactions as they are primarily responsible for the stiffness of the system. Thus we omit the long-range Coulomb or Lenard-Jones interaction potentials.

*2-body interactions.* The pair-wise bond potential  $V_{\text{bond}}$  depends on the distance  $r^{(i)} = |q^{(i+1)} - q^{(i)}|$ ,  $i = 1, \dots, N-1$  of the bonded particles  $(q^{(i)}, q^{(i+1)})$  in the linear chain. Typically this potential is assumed to be harmonic with the equilibrium distance  $r^{(i)} = r_0$ . In all simulations considered here the bonds are treated as rigid bonds of the constant length  $r_0 = 1$ . Thus defining the constrained DOFs

$$\xi_{\text{bond}}(q) = (r^{(1)}, \dots, r^{(N-1)}) = (1, \dots, 1).$$

*3-body interactions.* The interaction of three consecutive particles  $(q^{(i)}, q^{(i+1)}, q^{(i+2)})$  in the chain is defined by the potential  $V_{\text{angle}}(\theta^{(i)})$ ,  $i = 1, \dots, N-2$  that depends on the angle between the vectors  $(q^{(i+2)} - q^{(i+1)})$  and  $(q^{(i+1)} - q^{(i)})$ . In the simulated model we assume the bending angle potential

$$V_{\text{angle}}(\theta) = \frac{A_0}{2} \sin^2(\theta - \frac{\pi}{2}).$$

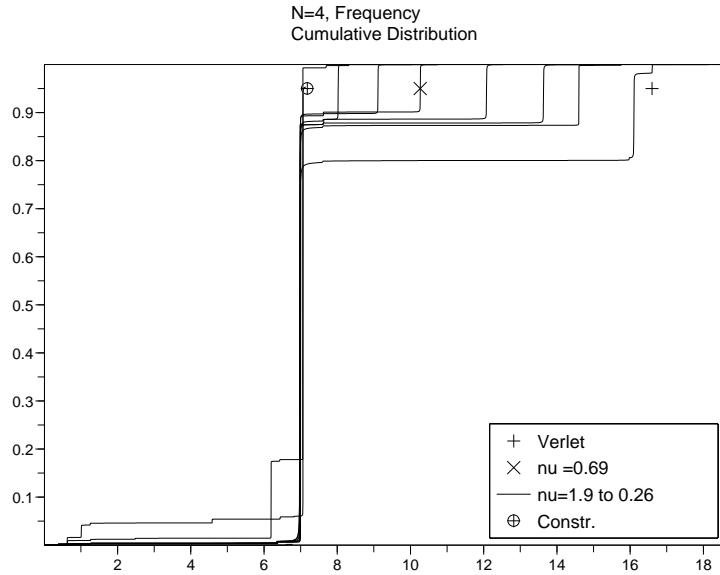


FIGURE 5.2. Frequency distribution of the end-to-end length oscillations of the butane Verlet dynamics and of the IMMP dynamics, cf. Figure 5.1. Note the two main frequencies (slow torsion and fast bond angles oscillations), and the slow components due to fluttering of resonances.

Depending on simulation the bending angles can be mass-penalized or directly constrained leading to the definition of penalized DOFs

$$\xi_{\text{angle}}(q) = (\theta^{(1)}, \dots, \theta^{(N-2)}) = \frac{z_a}{\nu} \quad \text{in the case of mass-penalization,}$$

$$\xi_{\text{angle}}(q) = (\theta^{(1)}, \dots, \theta^{(N-2)}) = 0 \quad \text{in the case of rigid constraints,}$$

where  $z_a \in \mathbb{R}^{N-2}$  are the auxiliary variables of the IMMP method.

*4-body interactions.* The last contribution  $V_{\text{torsion}}(\phi^{(i)})$  to the interaction potential depends on the torsion (dihedral) angles  $\phi^{(i)}$ ,  $i = 1, \dots, N - 3$  which are defined as angles between two planes  $\text{span}\{(q^{(i+3)} - q^{(i+2)}), (q^{(i+2)} - q^{(i+1)})\}$  and  $\text{span}\{(q^{(i+2)} - q^{(i+1)}), (q^{(i+1)} - q^{(i)})\}$ . We use a simple choice of the torsion potential

$$V_{\text{torsion}}(\phi) = -B_0 \cos(\phi).$$

The torsion DOFs are mass-penalized with the auxiliary variable  $z_b \in \mathbb{R}^{N-3}$

$$\xi_{\text{torsion}}(q) = (\phi^{(1)}, \dots, \phi^{(N-3)}) = \frac{z_b}{\nu}.$$

*System of units.* For the purpose of computational tests we have chosen the system of units in which the mass of united atoms is  $m^{(1)} = \dots = m^{(N)} = m_0 = 1$ , the equilibrium bond lengths are  $r^{(1)} = \dots = r^{(N-1)} = r_0 = 1$  and the inverse temperature  $\beta = 1/kT = 1$  at the ambient temperature  $T = 300 K$ . The time step  $\delta t = 0.01$  in these units corresponds to the physical time step  $\delta \tilde{t} = \sqrt{\beta m_0 r_0} = 3 \text{ fs}$ . Following physically relevant parameters (see, for example, [33, 30]) with an artificially slightly stiffer bending angle potential leads to the angle potential constant  $A_0 = 500$  and the torsion potential constant  $B_0 = 20$ .

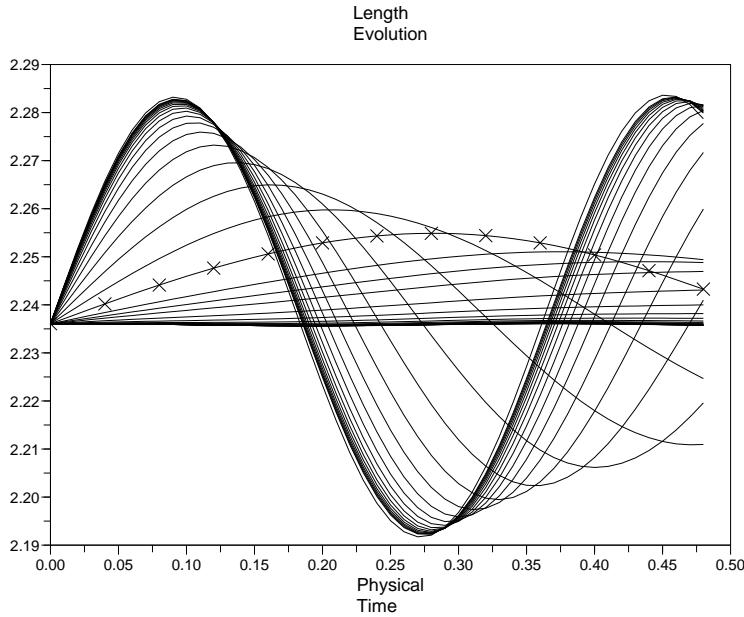


FIGURE 5.3. Interpolation of short trajectories of the butane length with IMMP dynamics, from constrained to Verlet dynamics. The marked trajectory corresponds to  $\nu = 5.5$ . The Verlet dynamics is the oscillating limit.

Applying direct constraints to bonds or bending angles in the presented formulation of the IMMP method consists in replacing the auxiliary variables  $z$  associated with fully constrained DOFs by 0. As described in Proposition 3.7 on the large penalty limit (see also Appendix A and Section 6 for the limit of stiff distributions), the associated dynamics is the standard ‘‘rigid’’ constrained dynamics. Then for sampling, the Fixman geometric corrector (1.4) is computed and accounts for the geometric bias of all the constrained DOFs including bonds. Note that the physical relevance of the geometric bias due to bonds is ambiguous since the quantum resolution may need to be introduced for the bond interactions. Such a change can be easily incorporated by computing the Fixman corrector potential associated with the bonds only and subtracting it from the Fixman corrector potential related to all the constrained DOFs.

**5.2. Efficiency criteria.** The efficiency of the IMMP method as compared to the Verlet/Leapfrog integrator is quantified from two different viewpoints.

*Dynamics.* To compare computational efficiency of numerical schemes, a notion of critical time step has to be introduced. The critical time step  $\delta t_c^{\text{dyn}}$  is defined implicitly through the formula

$$\int \frac{\beta}{3N - N_c} [H_{n+1} - H_n]^+ \mu_\nu(p_\nu^n, q^n) = \alpha, \quad (5.1)$$

where  $N_c$  denotes the number of rigidly constrained DOFs,  $[.]^+$  is the positive part,  $H_n$  is the energy of the numerical scheme at the  $n$ -th step, and  $\mu_\nu$  is the Boltzmann distribution associated with the Hamiltonian at hand as defined in (3.8). The quantity  $\alpha$  is a prescribed typical non-dimensional error of the energy per degrees of freedom as compared to the inverse temperature  $\beta$ . At least when  $\alpha$  is small, this defines uniquely  $\delta t_c^{\text{dyn}}$ . Thus for a given  $\alpha$ , the larger  $\delta t_c^{\text{dyn}}$  is, the less costly the method is in terms of force evaluations per integration time step.

The critical time step  $\delta t_c^{\text{dyn}}$  is compared for the IMMP method and the Verlet/Leapfrog scheme. To achieve a fair comparison, it has then to be checked that the IMMP penalty, which introduces a dynamical

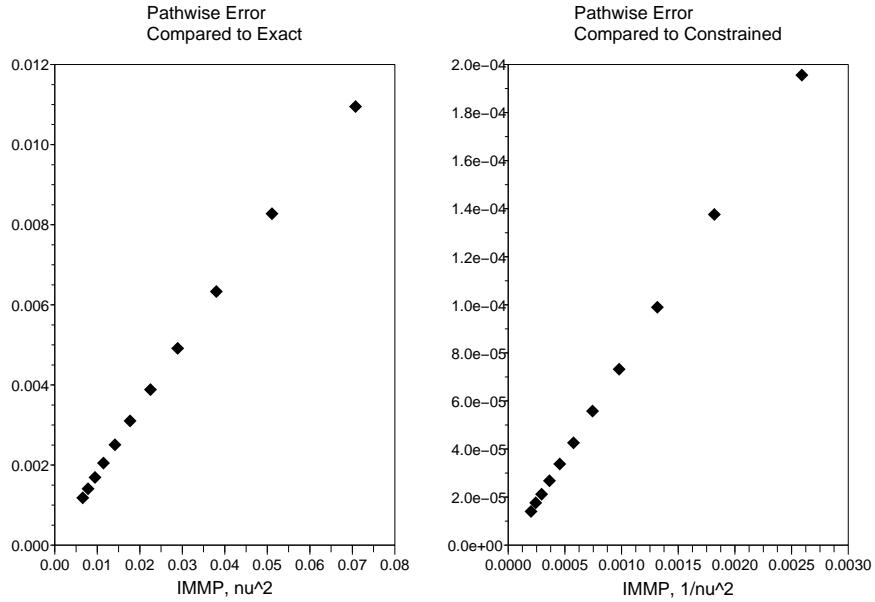


FIGURE 5.4. Convergence in terms of the  $\ell^2$  pathwise error of the short time butane length IMMP dynamics compared to Verlet dynamics and constrained dynamics, with orders of convergence  $\mathcal{O}(\nu^2)$  and  $\mathcal{O}(\nu^{-2})$ .

modification, does not modify the relevant slow frequencies and slow varying components of the system on large time intervals. This can be done by comparing the distortion of the frequency distribution of a long trajectory. In the numerical tests we analyze for  $t \in [0, T]$  the trajectory of the end-to-end molecule length  $t \mapsto L(t)$ , and we define the frequency density as the normalized square modulus of its Fourier transform  $d(\omega) = |\hat{L}(\omega)|^2/Z_f$  where  $Z_f$  is the normalization making  $d(\omega)$  a probability density function, and  $\hat{L}(\omega) = \int_0^T L(t) \exp(i\omega t) dt$ . We plot the cumulative distribution function

$$D(\omega) = \int_0^\omega d(\omega') d\omega'.$$

Hence, in the figures dominant frequencies correspond to jumps in the associated cumulative distribution. Of course, this comparison remains still largely qualitative.

*Sampling.* When using a Metropolis correction step in the numerical simulation, as in Scheme 4.2, the critical time step is correlated to the rejection rate of the Metropolis step, since the larger the former is, the more rejections will occur. Thus for sampling methods, the time step  $\delta t_c^{\text{samp}}$  is tuned in order to achieve a given rate of rejection  $\rho$  satisfying

$$\int \exp(-\beta[H_{n+1} - H_n]^+) \mu_\nu(p_\nu^n, q^n) = 1 - \rho, \quad (5.2)$$

where  $\exp(-\beta[H_{n+1} - H_n]^+)$  is the Metropolis weight that appears in Generalized Hybrid Monte-Carlo methods, as defined in Scheme 4.2. Then, computational efficiency is defined by comparing  $\delta t_c^{\text{samp}}$  with the mixing time *in terms of physical time* of the resulting Markov Chain. Equivalently, and more directly, the results are presented as the mixing time *in terms of iteration steps*. Such comparison requires choosing an appropriate notion of the mixing time. At least when the momenta are overdamped (i.e.,  $\gamma \rightarrow +\infty$ ), the chain becomes reversible, and the mixing time can be rigorously defined using the spectral gap of the

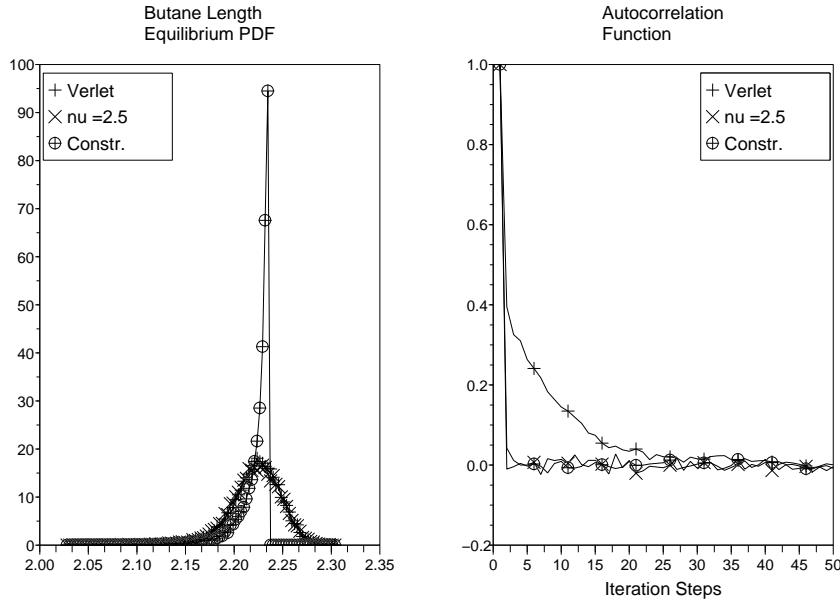


FIGURE 5.5. Left: Equilibrium PDF of the end-to-end length of the butane molecule with the GHMC scheme, using Verlet, IMMP (penalty  $\nu$ ), and constrained/RATTLE integrators. Note that the constrained integrator does not sample the correct measure. Right: The autocorrelation function in terms of iteration steps for the GHMC scheme comparing the IMMP and the Verlet integrator. The decrease in the  $\ell^2$ -decorrelation time (5.3) is by the factor 1.8.

underlying Markov kernel. In the present work, the decorrelation time of relevant observables is used. Assuming the initial state  $q_0$  of the system is at equilibrium, the normalized autocorrelation function associated with a given position observable  $q \mapsto \phi(q)$  is

$$C_n(\phi) := \frac{\mathbb{E}[\phi(q_n)\phi(q_0)]}{\mathbb{E}[\phi(q_0)^2]},$$

and it can be computed in large time simulations using path averages. Then  $\ell^2$ -decorrelation time  $n_{\text{corr}}$ , calculated in terms of iteration steps, is given by the formula

$$n_{\text{corr}} = 2 \sum_{n=0}^{+\infty} C_n(\phi)^2. \quad (5.3)$$

This quantity corresponds to the approximate number of steps of the chain for the observable  $\phi$  to decorrelate from its past values.

**5.3. Numerical results.** For comparisons we choose an often studied observable defined as the end-to-end distance of the alkane chain.

I. THE BUTANE MODEL ( $N = 4$ ). The bonds between atoms are rigidly constrained, and bond angles are treated using the IMMP method with Scheme 4.1.

*Dynamical behavior.* The dynamics of the butane model is studied using deterministic dynamics with a prescribed initial energy. As shown in Figure 5.1, the IMMP dynamics with different penalty  $\nu$  yield an interpolation from the mixed torsion/bond angles oscillations of the exact dynamics, to the simple torsion oscillation of the constrained dynamics. Depending on the frequency introduced by the mass penalization, some fluttering resonance can be observed in Figure 5.1 between the torsion and the bond angles.

Behavior observed in Figure 5.1 is related to the frequency density of the end-to-end length oscillations which are depicted in Figure 5.2. The bond angles oscillation frequency appears to slow down with the IMMP penalization and is eventually setting to a single frequency of the constrained dynamics. Note the slow modes introduced by the resonances.

The dynamical interpolation of the IMMP from the exact dynamics to the constrained dynamics is demonstrated on short time trajectories in Figure 5.3. The associated convergence orders,  $\mathcal{O}(\nu^2)$  and  $\mathcal{O}(\nu^{-2})$ , respectively, are captured in Figure 5.4. The time step stability is studied in Table 5.1. The IMMP method enables an increase of the critical time step of the Verlet scheme, following the interpolation property.

*Sampling behavior.* Exact sampling of the equilibrium distribution on a very large time scale, whatever the value of the IMMP penalization, is shown in Figure 5.5. The distribution of the butane length for constrained bond angles is clearly distorted. The mixing time to equilibrium is also studied. The autocorrelation function of the length evolution in terms of iteration steps is plotted in Figure 5.5, and the faster convergence of the IMMP method is demonstrated. The  $\ell^2$ -decorrelation time (5.3) is decreased by the factor 1.8 using the IMMP method.

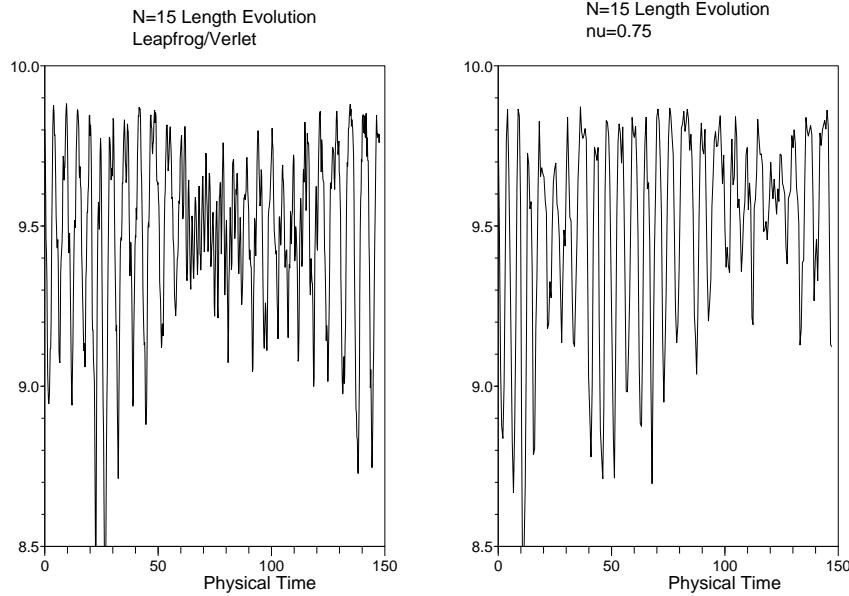


FIGURE 5.6. The trajectory of the  $N = 15$ -alkane dynamics for the Verlet scheme and the IMMP scheme. Note that the IMMP penalty does not modify substantially the low frequencies/slowly varying components. The frequency analysis is presented in Figure 5.7

II. THE ALKANE MODEL ( $N = 5, \dots, 20$ ). In this test the bonds and bending angles between atoms are rigidly constrained, and torsion angles are treated using the IMMP method. Note that the rigid constraints on torsions would lead to a rigid molecule, losing completely the evolution of the molecule length. The IMMP penalty is increased with the system size using the linear scaling  $\nu = \bar{\nu}N$ . In Section 7 we present a mathematical justification of this scaling for linear systems. In principle, introducing inertia in the torsion angles adds weights on the diagonal of the mass-tensor in the internal coordinates of the molecule. This can reduce the lowest eigenvalues of the mass-tensor, and can explain the reduction of the multi-scale nature of the system that are observed in the simulations below.

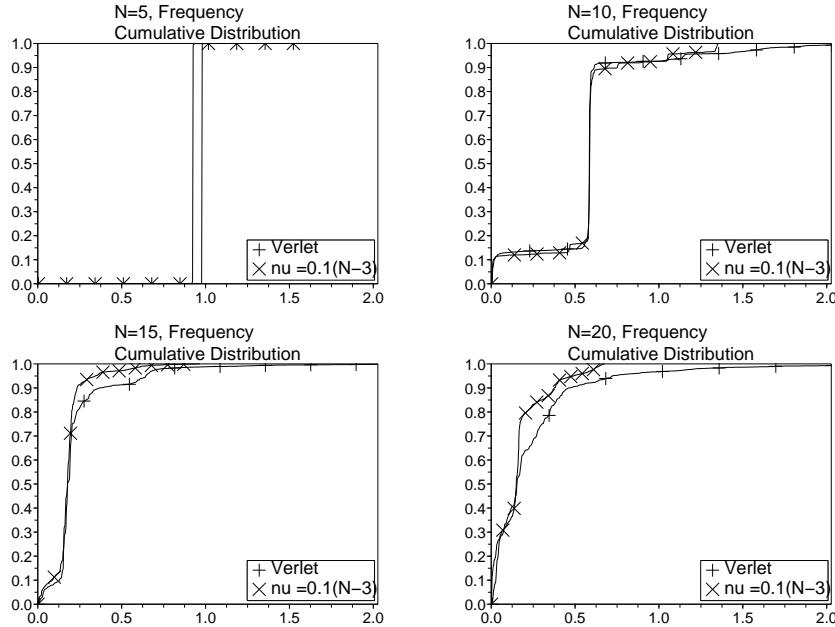


FIGURE 5.7. Spectral densities of the end-to-end length of the alkane for  $N = 5, 10, 15, 20$ . Note that the IMMP penalty does not modify substantially the low frequencies/slowly varying components.

*Dynamical behavior.* The frequency of oscillations of the alkane length in Figure 5.6, and in the spectral analysis in Figure 5.7 are not substantially modified by the IMMP penalization. One can observe a small group of fast oscillations in the middle of the Verlet dynamics plot in Figure 5.6 which is not present in the IMMP case. This translates in the top of the spectral plots in Figure 5.7 where a cut-off of the fastest oscillatory scales for the IMMP case occurs.

The critical time steps  $\delta t_c^{\text{dyn}}$  with respect to the system size are depicted in Figure 5.8. The gain in time stepping increases more than linearly with the system size  $N$ . This behavior, however, depends on the initial scaling  $\nu = \bar{\nu}N$  of the IMMP penalty.

*Sampling behavior.* Exact sampling of the equilibrium distribution on very large time scales (whatever the value of the IMMP penalization) is shown for  $N = 10$  in Figure 5.9, with the auto-correlation of the latter observable with the gain in mixing time of the IMMP dynamics. The precise ratio of the  $\ell^2$ -decorrelation time between the IMMP integrator and the Verlet one is given in Figure 5.11 for different system sizes. It increases again more than linearly in  $N$ , in fact exponentially for this particular system. The associated critical time steps  $\delta t_c^{\text{sampl}}$  are in Figure 5.10. We observe that the critical time step increases with large  $N$  which demonstrates that in the present case, the IMMP method heals the decrease of the Metropolis rejection rate of for large systems (see also [24]).

*Conclusions.* The presented numerical studies demonstrate that for integrating the dynamics, the IMMP allows for relaxation of the time-step stability restrictions. In the case of sampling methods the IMMP method decreases the decorrelation time measured in terms of Monte-Carlo iteration steps leading to more efficient sampling algorithm. In both cases the improvements are increasing with system size  $N$ .

**6. The infinite stiffness limit.** Throughout this section, one introduces a potential function with an explicit dependence with respect to the fast variables  $(q, z) \mapsto U(q, z)$  together with a stiffness parameter

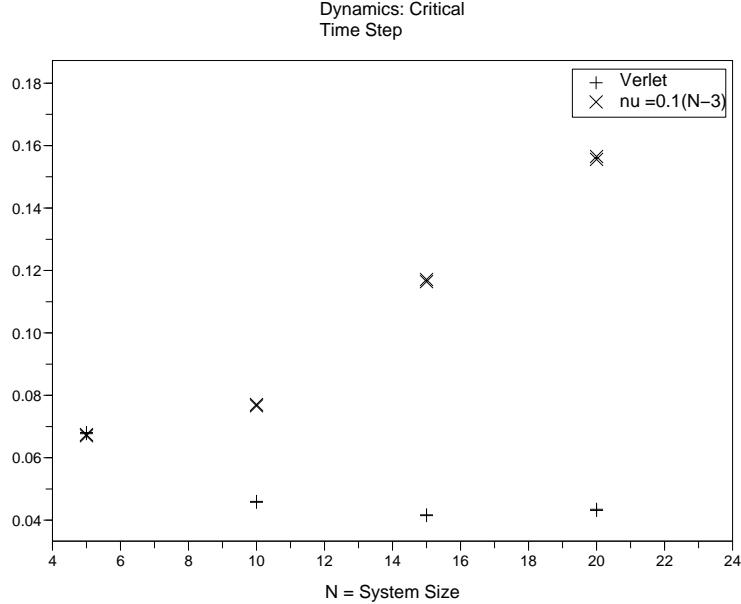


FIGURE 5.8. Critical time steps, with error bars,  $\delta t_c^{\text{dyn}}$  of the Verlet and IMMP dynamics with respect to the system size.

$\epsilon$ . The potential energy  $V$  can then be written in the form

$$V(q) = U(q, \frac{\xi(q)}{\epsilon}), \quad \text{with a confining assumption } \inf_{q \in \mathbb{R}^d} U(q, z) \geq K(z),$$

where  $K(z) > c \log |z|$  as  $|z| \rightarrow \infty$ . The functions  $\xi$  are then indeed ‘‘fast’’ degrees of freedom (fDOFs) in the limit  $\epsilon \rightarrow 0$ , the system being confined to the slow sub-manifold  $\mathcal{M}_0 = \{q \mid \xi(q) = 0\}$ .

We prove, that under appropriate scaling of the mass penalty  $\nu_\epsilon = \frac{\bar{\nu}}{\epsilon}$ , the IMMP method is asymptotically stable in the stiff limit, converging towards standard effective dynamics on the slow manifold  $\mathcal{M}_0$ .

### 6.1. Thermostatted stiff systems.

The canonical distribution becomes

$$\mu_\epsilon(dp dq) = \frac{1}{Z_\epsilon} e^{-\beta(\frac{1}{2}p^T M^{-1}p + U(q, \frac{\xi(q)}{\epsilon}))} dp dq. \quad (6.1)$$

In the infinite stiffness limit ( $\epsilon \rightarrow 0$ ) the measure concentrates on the slow manifold  $\mathcal{M}_0$ . The limit is computed using the co-area formula (see Appendix A for relevant definitions of surface measures). In order to characterize the limiting measure we introduce the effective potential

$$V_{\text{eff}}(q) = -\frac{1}{\beta} \ln \int e^{-\beta U(q, z)} dz. \quad (6.2)$$

LEMMA 6.1. *In the infinite stiffness limit ( $\epsilon \rightarrow 0$ ), the highly oscillatory canonical distribution (6.1) converges  $\mu_\epsilon \rightarrow \mu_0$  (in distribution) towards  $\mu_0(dp dq)$ , which is supported on  $\mathcal{M}_0$ , and defined as*

$$\mu_0(dp dq) = \frac{1}{Z_0} e^{-\beta(\frac{1}{2}p^T M^{-1}p + V_{\text{eff}}(q))} dp \delta_{\xi(q)=0}(dq). \quad (6.3)$$

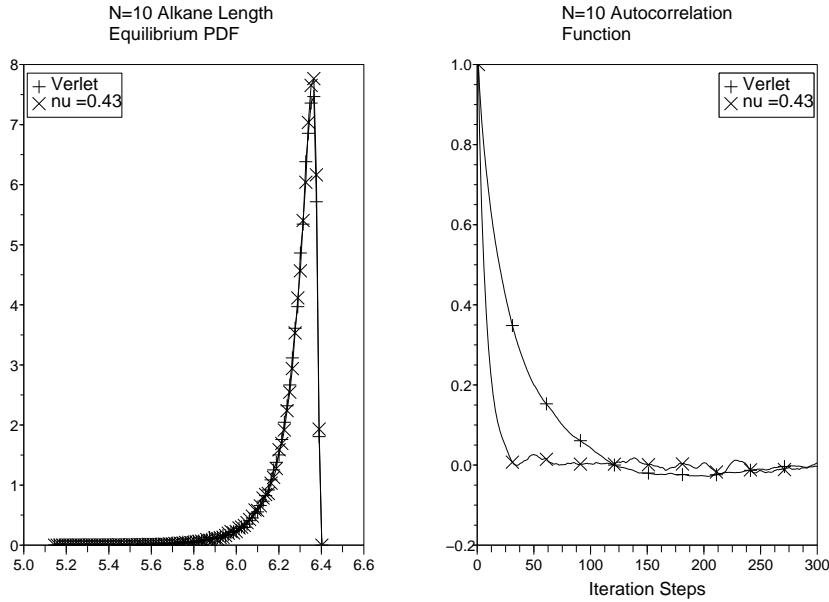


FIGURE 5.9. Left: Equilibrium PDF of the end-to-end length of the alkane chain with the GHMC scheme, using Verlet and IMMP integrator (penalty  $\nu$ ). Right: The autocorrelation function in terms of iteration steps for the GHMC scheme comparing the IMMP and the Verlet integrator. The system size is  $N = 10$

Its marginal distribution in position is given, up to the normalization, by

$$e^{-\beta V_{\text{eff}}(q)} \delta_{\xi(q)=0}(dq). \quad (6.4)$$

*Proof.* It is sufficient to consider distributions in the position variable  $q$  only. Let  $\mathcal{U}^\delta$  be a  $\delta$ -neighborhood of  $\mathcal{M}_0$  where  $dq = \epsilon^n \delta_{\xi(q)=\epsilon z}(dq) dz$ . We construct a decomposition  $\varphi = \varphi_1 + \varphi_2$  of continuous bounded observables such that  $\text{supp } \varphi_1 \subset \mathcal{U}^\delta$  and  $\text{supp } \varphi_2 \cap \mathcal{U}^{\delta/2} = \emptyset$ . Using the confining property of  $U(q, \cdot)$  we obtain

$$\int \varphi(q) e^{-\beta U(q, \frac{\xi(q)}{\epsilon})} dq = \epsilon^n \int \varphi_1(q) e^{-\beta U(q, z)} \delta_{\xi(q)=\epsilon z}(dq) dz + \mathcal{O}(e^{-\beta K(\delta/2\epsilon)}).$$

By continuity of  $\epsilon \mapsto \int \varphi_1(q) e^{-\beta U(q, z)} \delta_{\xi(q)=\epsilon z}(dq)$  and by the dominated convergence theorem

$$\int \varphi_1(q) e^{-\beta U(q, z)} \delta_{\xi(q)=\epsilon z}(dq) dz \rightarrow \int \varphi_1(q) e^{-\beta V_{\text{eff}}(q)} \delta_{\xi(q)=0}(dq) = \int \varphi(q) e^{-\beta V_{\text{eff}}(q)} \delta_{\xi(q)=0}(dq),$$

and the result follows after normalization.  $\square$

The infinite stiffness limit ( $\epsilon \rightarrow 0$ ) of highly oscillatory dynamics has been studied in a series of papers [37, 39, 27, 5, 34, 35]. The limiting dynamics can be fully characterized in special cases. For example, when the highly oscillatory potential is linear and non-resonant (at least almost everywhere on the trajectory, see [39]), it can be described through adiabatic effective potentials. See also [10, 6] for a recent work on some related numerical issues. However, when the system is thermostatted, one can postulate an ‘ad hoc’ effective dynamics ([35]) exhibiting the appropriate limiting canonical distribution given by (6.4). Such dynamics can be obtained by constraining the system to the slow manifold  $\mathcal{M}_0$ , and adding a correcting

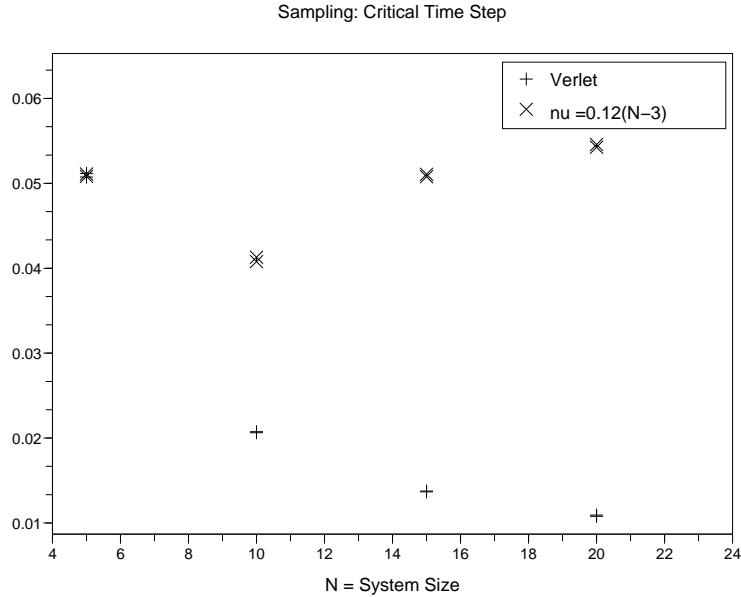


FIGURE 5.10. Critical time steps, with error bars,  $\delta t_c^{\text{sampled}}$  of the Verlet and IMMP dynamics with respect to the system size. Note that the IMMP penalty heals the degeneracy of the rejection rate for large systems.

entropic potential, sometimes called Fixman corrector from [19], which is due to the geometry of  $\mathcal{M}_0$ , and is given by

$$V_{\text{fix}}(q) = \frac{1}{2\beta} \ln (\det G(q)) , \quad (6.5)$$

where  $G(q)$  is the  $n \times n$  Gram matrix defined in (3.2).

In general, since the effective potential (6.2) is not explicit, one may need to couple the system with virtual fast degrees of freedom to enforce the appropriate effective dynamics associated with (6.2). The resulting extended Hamiltonian is then defined on the state space  $T^*(\mathcal{M}_0 \times \mathbb{R}^n)$  (the cotangent bundle) and is given by

$$\begin{cases} H_{\text{eff}}(p, p_z, q, z) = \frac{1}{2} p^T M^{-1} p + \frac{1}{2} p_z^T M_z^{-1} p_z + U(q, z) + V_{\text{fix}}(q) \\ \xi(q) = 0 . \end{cases} \quad (6.6) \quad (C)$$

DEFINITION 6.2 (Effective Langevin process with constraints). *The constrained Langevin process associated with Hamiltonian (6.6) is defined by the following stochastic differential equations*

$$\begin{cases} \dot{q} = M^{-1} p \\ \dot{z} = M_z^{-1} p_z \\ \dot{p} = -\nabla_1 U(q, z) - \nabla_q V_{\text{fix}}(q) - \gamma \dot{q} + \sigma \dot{W} - \nabla_2 \xi \dot{\lambda} \\ \dot{p}_z = -\nabla_2 U(q, z) - \gamma_z \dot{z} + \sigma_z \dot{W}_z \\ \xi(q) = 0 , \end{cases} \quad (6.7) \quad (C)$$

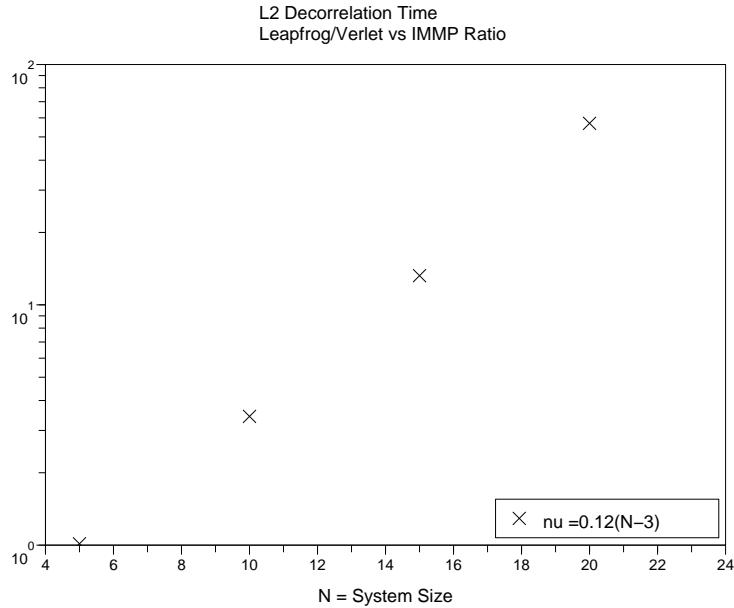


FIGURE 5.11. Comparison the  $\ell^2$ -decorrelation time of the end-to-end alkane length in terms of Monte-Carlo iteration steps. The ratio between the case of the Verlet integration and the case of the IMMP integration is depicted. Note that the vertical axis is in the logarithmic scale and an exponential gain occurs. Error bars are plotted, but do not appear at this scale.

where  $\nabla_1$  and  $\nabla_2$  are respectively derivatives with respect to the first and second variable of the function  $U(q, z)$ ,  $\dot{W}$  (resp.  $\dot{W}_z$ ) is the standard multi-dimensional white noise,  $\gamma$  (resp.  $\gamma_z$ ) a  $d \times d$  (resp.  $n \times n$ ) symmetric positive semi-definite dissipation matrix,  $\sigma$  (resp.  $\sigma_z$ ) is the fluctuation matrix satisfying  $\sigma\sigma^T = \frac{2}{\beta}\gamma$  (resp.  $\sigma_z\sigma_z^T = \frac{2}{\beta}\gamma_z$ ). The processes  $\lambda \in \mathbb{R}^n$  are Lagrange multipliers associated with the constraints (C) and adapted with respect to the white noise.

We formulate reversibility of this process as a separate lemma.

LEMMA 6.3. *The process defined in (6.7) is reversible with respect to the associated canonical distribution whose marginal distribution in  $(q, p)$  variables is*

$$\mu_{\text{eff}}(dp dq) = \frac{1}{Z_{\text{eff}}} e^{-\beta(\frac{1}{2}p^T M^{-1}p + V_{\text{eff}}(q) + V_{\text{fix}}(q))} \sigma_{T^* \mathcal{M}_0}(dp dq) \quad (6.8)$$

with the  $q$ -marginal

$$e^{-\beta V_{\text{eff}}(q)} \delta_{\xi(q)=0}(dq).$$

When  $\gamma$  and  $\gamma_z$  are strictly positive definite, the process is ergodic.

*Proof.* The process (6.7) is a Langevin process with mechanical constraints, exhibiting reversibility properties with respect to the associated Boltzmann canonical measure (see the summary in Appendix B). Then the  $q$ -marginal is obtained by remarking that the integration of any function of  $\frac{1}{2}p^T M^{-1}p + \frac{1}{2}p_z^T M_z^{-1}p_z$  with respect to  $dp_z \sigma_{T_q^* \mathcal{M}_0}(dp)$  results in a constant independent of  $q$ .  $\square$

The properties of thermostatted highly oscillatory systems are summarized in Table 6.1.

	Finite stiffness $\epsilon > 0$	Infinite stiffness limit $\epsilon \rightarrow 0$	Infinite stiffness $\epsilon = 0$
Dynamics	Highly oscillatory + fluct./diss.	Adiabatic (if non-resonant) + non-Markov fluct./diss.	Effective with constraints + fluct./diss.
Statistics	Canonical	Positions on $\mathcal{M}_0$ , free velocities.	Canonical on $T^*\mathcal{M}_0$ , geometric corrector.
Numerics	Leapfrog/Verlet + fluct./diss.	Time-step restrictions ( $\delta t = o(\epsilon)$ )	Leapfrog/Verlet with constraints + fluct./diss.

TABLE 6.1

Stiff Hamiltonian systems and associated commonly used numerical methods ( $\mathcal{M}_0$  denotes the slow manifold). Two different schemes are required for the stiff system and its effective Markovian approximation.

**6.2. Stability of the IMMP dynamics.** We assume that the mass-matrix penalty parameter  $\nu \equiv \nu_\epsilon$  grows to infinity in such a way that  $\lim_{\epsilon \rightarrow 0} \epsilon \nu_\epsilon = \bar{\nu}$ .

The original Hamiltonian with the stiffness parameter is expressed explicitly as

$$H_\epsilon(p, q) = \frac{1}{2} p^T M^{-1} p + U(q, \frac{\xi(q)}{\epsilon}), \quad (6.9)$$

and including the mass-matrix penalization one gets

$$H_{\nu_\epsilon}(p_{\nu_\epsilon}, q) = \frac{1}{2} p_{\nu_\epsilon}^T M_{\nu_\epsilon}^{-1} p_{\nu_\epsilon} + U(q, \frac{\xi(q)}{\epsilon}) + V_{\text{fix}, \nu_\epsilon}(q), \quad (6.10)$$

or in its implicit formulation

$$\begin{cases} H_{\text{IMMP}}(q, z, p, p_z) = \frac{1}{2} p^T M^{-1} p + \frac{1}{2} p_z^T M_z^{-1} p_z + U(q, \frac{z}{\nu_\epsilon \epsilon}) + V_{\text{fix}, \nu_\epsilon}(q), \\ \xi(q) = \frac{1}{\nu_\epsilon} z. \end{cases} \quad (C_{\nu_\epsilon}) \quad (6.11)$$

One immediately sees that  $H_{\text{IMMP}}$  is non-singular when  $\epsilon \rightarrow 0$  and converges to the effective Hamiltonian on the slow manifold,

$$\begin{cases} H_{\text{eff}, \bar{\nu}}(q, z, p, p_z) = \frac{1}{2} p^T M^{-1} p + \frac{1}{2} p_z^T M_z^{-1} p_z + U(q, \frac{z}{\bar{\nu}}) + V_{\text{fix}}(q) \\ \xi(q) = 0. \end{cases} \quad (C). \quad (6.12)$$

The expression (6.11) represents a minor generalization of  $H_{\text{eff}}$  in (6.6), but it leads to the same canonical marginal distribution  $\mu_{\text{eff}}(dp dq)$  in  $(p, q)$  variables as given by (6.8). The continuity in  $\epsilon$  of  $H_{\text{IMMP}}$  implies stability of the associated dynamics and their numerical integrators. We first derive the limits of the original and penalized canonical distribution.

**PROPOSITION 6.4** (Limits of canonical distributions). *Consider the canonical distributions  $\mu_{\nu_\epsilon}(dp dq)$  associated with the mass penalized Hamiltonian (6.10), but considered with respect to the variables  $(p \equiv MM_{\nu_\epsilon}^{-1}p_{\nu_\epsilon}, q)$ . In the sense of weak convergence of measures we have  $\mu_{\nu_\epsilon} \rightarrow \mu_{\text{eff}}$  as  $\epsilon \rightarrow 0$  with  $\mu_{\text{eff}}$  defined by (6.8).*

*Proof.* The first convergence is proved in Lemma 6.1. For the second one, the following notation will be used

$$\delta_{q, \epsilon z}(dq) = \delta_{\xi(q)=\epsilon z}(dq), \quad \text{and} \quad \delta_{p, \epsilon p_z}(dp) = \delta_{p^T M^{-1} \nabla \xi(q)=\epsilon M_z^{-1} p_z}(dp).$$

To prove the convergence towards  $\mu_{\text{eff}}$  we consider a  $\delta$ -neighborhood  $\mathcal{U}^\delta$  of  $\mathcal{M}_0$  where

$$dp \, dq = \frac{\epsilon^{2n}}{\det M_z} \delta_{q, \epsilon z}(dq) \, dz \, \delta_{p, \epsilon p_z}(dp) \, dp_z,$$

and a decomposition of the bounded observable (in  $(p, q)$  variables)  $\varphi = \varphi_1 + \varphi_2$  such that  $\text{supp } \varphi_1 \subset \mathcal{U}^\delta$  and  $\text{supp } \varphi_2 \cap \mathcal{U}^{\delta/2} = \emptyset$ . Thus, keeping in mind that  $p_{\nu_\epsilon} = M_{\nu_\epsilon} M^{-1} p$ , and using the confining property of the potential  $U(q, \cdot)$  we obtain

$$\int \varphi(p, q) e^{-\beta H_{\nu_\epsilon}} dp_{\nu_\epsilon} \, dq = \int \varphi_1(p, q) e^{-\beta H_{\nu_\epsilon}} dp_{\nu_\epsilon} \, dq + \mathcal{O}(e^{-\beta K(\delta/\epsilon)}) \equiv I_\epsilon + \mathcal{O}(e^{-\beta K(\delta/2\epsilon)}). \quad (6.13)$$

Applying the change of variables  $p_{\nu_\epsilon} = M_{\nu_\epsilon} M^{-1} p$  yields

$$dp_{\nu_\epsilon} = \det(M_{\nu_\epsilon} M^{-1}) \, dp = \nu_\epsilon^{2n} \det M_z \det(G + \frac{1}{\nu_\epsilon^2} M_z^{-1}) \, dp,$$

and setting  $\epsilon M_z^{-1} p_z = \nabla_q \xi M^{-1} p$  and  $\epsilon z = \xi(q)$  we get

$$H_{\nu_\epsilon}(p_{\nu_\epsilon}, q) = \frac{1}{2} p^T M^{-1} p + \nu_\epsilon^2 \epsilon^2 p_z^T M_z^{-1} p_z + U(q, \frac{z}{\nu_\epsilon \epsilon}) + V_{\text{fix}, \nu_\epsilon}(q) = H_{\text{IMMP}}(q, z, p, p_z).$$

Thus substituting back to (6.13) we obtain

$$I_\epsilon = (\nu_\epsilon \epsilon)^{2n} \int \varphi_1 e^{-\beta H_{\text{IMMP}}(q, z, p, p_z)} \det(G + \frac{1}{\nu_\epsilon^2} M_z^{-1}) \delta_{p, \epsilon p_z}(dp) \, dp_z \delta_{q, \epsilon z}(dq) \, dz,$$

and thus

$$I_\epsilon \xrightarrow{\epsilon \rightarrow 0} \bar{\nu}^{2n} \int \varphi_1 e^{-\beta H_{\text{eff}, \bar{\nu}}(q, z, p, p_z)} \det(G) \delta_{p, \epsilon p_z}(dp) \, dp_z \delta_{\xi(q)=0}(dq) \, dz.$$

Using the co-area formula we obtain

$$\det(G) \delta_{\nabla \xi(q) M^{-1} p = 0}(dp) \delta_{\xi(q)=0}(dq) = \sigma_{T^* \mathcal{M}_0}(dp \, dq),$$

which leads to the final result after integration of the  $(p_z, z)$  variables and normalization.  $\square$

**REMARK 6.5.** Due to the fast oscillations, the distribution of impulses in the limiting distribution  $\mu_0$  in (6.3) is uncorrelated, whereas after the mass-matrix penalization, the limiting distribution (3.8) has almost surely co-tangent impulses (i.e., satisfying the constraints  $\nabla_q \xi M^{-1} p = 0$ ). This explains the role of the corrected potential energy  $V_{\text{fix}}$  taking into account the curvature of  $\mathcal{M}_0$ .

In the next step we inspect the infinite stiffness asymptotic of the penalized dynamics.

**PROPOSITION 6.6** (Infinite stiffness limit). *When  $\epsilon \rightarrow 0$  with  $\nu \equiv \nu_\epsilon \sim \frac{\bar{\nu}}{\epsilon}$  and  $V(q, \xi(q)) = U(q, \frac{\xi(q)}{\epsilon})$ , the IMMP Langevin stochastic process (3.5) converges weakly towards the following coupled limiting processes with constraints*

$$\begin{cases} \dot{q} = M^{-1} p, \\ \dot{p} = -\nabla_1 U(q, \frac{z}{\bar{\nu}}) - \nabla_q V_{\text{fix}}(q) - \gamma \dot{q} + \sigma \dot{W} - \nabla_q \xi \dot{\lambda}, \\ \dot{\xi}(q) = 0, \\ \dot{z} = M_z^{-1} p_z, \\ \dot{p}_z = -\frac{1}{\bar{\nu}} \nabla_2 U(q, \frac{z}{\bar{\nu}}) - \gamma_z \dot{z} + \sigma_z \dot{W}_z. \end{cases} \quad (C) \quad (6.14)$$

where  $\nabla_1$  and  $\nabla_2$  are respectively derivatives with respect to the first and second variable of the function  $U(q, z)$ , and  $\{\lambda_t\}_{t \geq 0}$  are adapted stochastic processes defining the Lagrange multipliers associated with the constraints  $(C)$ .

The process  $\{q_t, p_t\}_{t \geq 0}$  defines an effective dynamics with constraints (Definition 6.2) for thermostatted highly oscillatory systems. It is reversible with respect to its stationary canonical distribution given by  $\mu_{\text{eff}}$  (6.8), and is ergodic when  $(\gamma, \gamma_z)$  are strictly positive definite.

*Proof.* The proof is similar to the proof of Proposition 3.7. Here we have

$$\nabla_q U = \nabla_1 U + \frac{1}{\epsilon} \nabla_q \xi \nabla_2^T U,$$

and (3.5) translates, up to a change of Lagrange multipliers, into

$$\begin{cases} \dot{q} = M^{-1} p \\ \dot{z} = M_z^{-1} p_z \\ \dot{p} = -\nabla_1 U - \nabla_q V_{\text{fix}, \nu}(q) - \gamma \dot{q} + \sigma \dot{W} - \nabla_q \xi \dot{\lambda} \\ \dot{p}_z = -\frac{1}{\nu_\epsilon \epsilon} \nabla_2 U - \gamma_z \dot{z} + \sigma_z \dot{W}_z + \frac{\dot{\lambda}}{\nu_\epsilon} \\ \xi(q) = \frac{z}{\nu_\epsilon}, \end{cases} \quad (C_{\nu_\epsilon}). \quad (6.15)$$

The rest follows the proof of Proposition 3.7.

□

REMARK 6.7. When  $\bar{\nu} \rightarrow +\infty$ , by a classical averaging argument (see, e.g., [27]), one can check that the limiting dynamics are the effective dynamics pointed out in [35]

$$\begin{cases} \dot{q} = M^{-1} p \\ \dot{p} = -\nabla_q U_{\text{eff}}(q) - \nabla_q V_{\text{fix}} - \gamma \dot{q} + \sigma \dot{W} - \nabla_q \xi \dot{\lambda} \\ \xi(q) = 0. \end{cases} \quad (C) \quad (6.16)$$

with the stationary canonical distribution (6.8).

**6.3. Stability of the IMMP integrator.** The numerical scheme (Scheme 4.1 proposed for the IMMP method (3.5)) is also stable in the limit of infinite stiffness  $\epsilon \rightarrow 0$ . Recall that we consider a reversible, measure preserving numerical flow  $\Phi_{\delta t}^{\nu_\epsilon}(p, p_z, q, z)$  associated with Hamiltonian (6.11)  $H_{\text{IMMP}}$  with constraints (modified potentials could similarly be considered).

PROPOSITION 6.8 (Asymptotic stability). *In the limit  $\epsilon \nu_\epsilon \rightarrow \bar{\nu}$ , the numerical flow  $\Phi_{\delta t}^{\nu_\epsilon}$  associated with the leapfrog/Verlet integrator with constraints for the IMMP Hamiltonian (6.11) converges towards the numerical flow  $\Phi_{\delta t}^{\bar{\nu}}$ , which is the leapfrog/Verlet integrator with geometric constraints associated with effective Hamiltonian (6.12) on the slow manifold.*

*Proof.* The statement is a direct consequence of the implicit function theorem and the continuity of the leapfrog integrator with constraints (4.1) with respect to the parameter  $\bar{\nu} = \epsilon \nu_\epsilon$ . Indeed, considering the shift of Lagrange multipliers  $\lambda \rightarrow \lambda + \frac{1}{\epsilon} \nabla_2 U$  and taking the limit  $\epsilon \rightarrow 0$  we obtain the appropriate

leapfrog scheme

$$\left\{ \begin{array}{l} p_{n+1/2} = p_n - \frac{\delta t}{2} \nabla_1 U(q_n, \frac{z_n}{\bar{\nu}}) - \nabla_q \xi(q_n) \lambda_{n+1/2} \\ p_{n+1/2}^z = p_n^z - \frac{\delta t}{2\bar{\nu}} \nabla_2 U(q_n, \frac{z_n}{\bar{\nu}}) \\ q_{n+1} = q_n + \delta t M^{-1} p_{n+1/2} \\ z_{n+1} = z_n + \delta t M_z^{-1} p_{n+1/2}^z \\ \xi(q_{n+1}) = 0 \end{array} \right. \quad (C_{1/2})$$

$$\left\{ \begin{array}{l} p_{n+1} = p_{n+1/2} - \frac{\delta t}{2} \nabla_1 U(q_{n+1}, \frac{z_{n+1}}{\bar{\nu}}) - \nabla_q \xi(q_{n+1}) \lambda_{n+1} \\ p_{n+1}^z = p_{n+1/2}^z - \frac{\delta t}{2\bar{\nu}} \nabla_2 U(q_{n+1}, \frac{z_{n+1}}{\bar{\nu}}) \\ \nabla_q \xi(q_{n+1}) M^{-1} p_{n+1} = 0. \end{array} \right. \quad (C_1)$$

□

By convergence of the Hamiltonian (6.11) to (6.12), similar asymptotic stability properties holds when a Metropolis step is introduced.

The results and properties discussed in this section are summarized in Table 6.2.

	Zero mass penalization $\nu = 0$	Positive mass-penalization $\epsilon, \nu > 0$	Infinite stiffness limit $\epsilon \rightarrow 0, \frac{\nu}{\epsilon} \rightarrow \bar{\nu}$
Dynamics	Highly oscillatory + fluct./diss.	IMMP + fluct./diss.	Effective with constraints + fluct./diss.
Statistics	Canonical	Canonical with correlated velocities	Canonical on $T^* \mathcal{M}_0$
Numerics		IMMP + fluct./diss.	

TABLE 6.2

The IMMP dynamics and the Verlet numerical integration are both asymptotically stable in the infinite stiffness regime if  $\frac{\nu}{\epsilon} \rightarrow \bar{\nu} < +\infty$ . If the mass-penalization vanishes ( $\nu = 0$ ) one recovers the original physical stiff system. The canonical distribution is always exact in the position variable. Notice that due to the penalized mass-matrix ( $\nu > 0$ ) the statistics have correlated velocities.

**7. Numerical analysis of a harmonic particle chain.** In this section we present rigorous analysis for a special case of the linear chain with *harmonic* interactions. The analysis supports scaling properties of the IMMP, with respect to the size of the chain, observed in numerical simulations of the general linear alkane chains. We consider the thermodynamic limit  $N \rightarrow +\infty$  where  $N$  is the size of the system. It is shown that the macroscopic dynamics of the IMMP method behaves continuously (uniformly with  $N$ , and in the  $L_2$  norm for the position profile) with respect to the re-scaled mass penalty parameter  $\bar{\nu}$ .

At the same time, the time-step stability of the IMMP numerical scheme (4.3) is compared with the standard Verlet scheme, and the critical time step is shown to be increased by a factor  $\bar{\nu}N$ .

From the spectral point of view, the IMMP method behaves in this linear case as a low-pass filter. This proves, in this simplified case, the ability of IMMP method to respect macroscopic dynamical equivalence, while saving computational time up to a factor of order  $\mathcal{O}(N)$ .

**7.1. Conservation of macroscopic dynamics.** The model we consider consists of a chain of particles which interact through the harmonic (quadratic) potential  $v_{\text{int}}(r) = r^2/2$ . Each particle is also individually submitted to a macroscopic confining exterior (quadratic) potential  $v_{\text{ext}}(r)$ . After converting

to the non-dimensional form the typical quantities involved in the model enable us to write a scaling at the mass-transport level where the dynamics of the chain is described by the Hamiltonian

$$H_N(q, p) = \frac{1}{2}p^T p + \sum_{i=1}^{N-1} v_{\text{int}}(\nabla_i^d q) + \sum_{i=1}^N v_{\text{ext}}(q_i), \quad (7.1)$$

and by a coupling with an exterior thermal bath at the re-scaled inverse temperature  $\beta_N = \beta N^{-1}$ . In the expression (7.1) the functions  $r \in \mathbb{R} \mapsto v_{\text{int}}(r) \in \mathbb{R}$  and  $q \mapsto v_{\text{ext}}(q) \in \mathbb{R}$  are the smooth interaction potential and the exterior potential, respectively. The linear operator  $\nabla^d : \mathbb{R}^N \rightarrow \mathbb{R}^{N-1}$ , having the components

$$\nabla_i^d q = \frac{q_{i+1} - q_i}{1/N}, \quad i = 1, \dots, N-1,$$

represents the discrete gradient associated to the chain with the Neumann boundary conditions. Its transpose operator is denoted  $(\nabla^d)^T : \mathbb{R}^{N-1} \rightarrow \mathbb{R}^N$ . The discrete Laplace operator is then defined as  $\Delta_d = -(\nabla^d)^T \nabla^d$ . The particles are represented by their re-scaled positions  $q = (q_1, \dots, q_N)$ , so that the typical position and deviation of  $q$  is formally of order 1 with respect to  $N$ . This can be seen by considering particles in the chain as indexed by  $x = \frac{i}{N} \in [0, 1]$ . We choose to work with such scaling in  $N$  that it prescribes the macroscopic timescale of the chain profile at order one with respect to  $N$ .

Following our general construction we obtain the mass-penalized Hamiltonian

$$H_{\nu_N}(p_{\nu_N}, q) = \frac{1}{2}p_{\nu_N}^T M_{\nu_N}^{-1} p_{\nu_N} + \sum_{i=1}^N v_{\text{int}}(\nabla_i^d q) + \sum_{i=1}^N v_{\text{ext}}(q_i). \quad (7.2)$$

We chose the penalizing matrix to be the identity matrix  $M_z = \text{Id}$ , hence the penalized mass-tensor becomes  $M_{\nu_N} = \text{Id} - \bar{\nu}^2 \Delta_d$ , and the fluctuation/dissipation tensor is taken proportional to the identity matrix. The system of stochastically perturbed equations of motions then becomes

$$\begin{cases} \dot{q} = (\text{Id} - \bar{\nu}^2 \Delta_d)^{-1} p \\ \dot{p} = \Delta_d q - v'_{\text{ext}}(q) - \gamma \dot{q} + \sigma \sqrt{N} \dot{W}, \end{cases} \quad (7.3)$$

with fluctuation/dissipation identity  $\sigma^2 = 2\beta^{-1}\gamma$ . The associated canonical equilibrium distribution is then given by the re-scaled inverse temperature  $\beta_N = \beta N^{-1}$ .

In order to treat the limit  $N \rightarrow \infty$ , we introduce the  $\ell_2$ -norm in the position space

$$\|q\|_{\ell_2}^2 := \frac{1}{N} \sum_{i=1}^N q_i^2 = \frac{1}{N} q^T q,$$

as well as the  $h_{-1}$ -norm in the momentum space

$$\|p\|_{h_{-1}}^2 = \left\| \left( -\Delta_d \right)^{-1/2} \left( p - \frac{1}{N} \sum_{i=1}^N p_i \right) \right\|_{\ell_2}^2 + \left( \frac{1}{N} \sum_{i=1}^N p_i \right)^2.$$

In the above expression,  $\frac{1}{N} \sum_{i=1}^N p_i$  can be seen as the orthogonal projection in  $\ell^2$  on the one dimensional kernel of the Neumann discrete Laplacian  $\Delta_d$ . The quadratic form  $\|q\|_0^2 + \|p\|_{-1}^2$  endows the phase-space with a Hilbert space structure.

**PROPOSITION 7.1** (Convergence of the macroscopic dynamics). *Assume that the exterior potential  $v_{\text{ext}}$  is bounded and that its derivative satisfies the Lipschitz condition*

$$\|v'_{\text{ext}}(q_2) - v'_{\text{ext}}(q_1)\|_{h_{-1}} \leq L_v \|q_2 - q_1\|_{\ell_2},$$

where  $L_v$  is independent of  $N$ . For any  $T > 0$ , let  $t \mapsto (p^{\bar{\nu}}(t), q^{\bar{\nu}}(t))$  be the solution, for  $t \in [0, T]$ , of the evolution equation (7.3) with the initial condition

$$(p^{\bar{\nu}}(0), q^{\bar{\nu}}(0)) = (M_{\nu_N}^{-1/2} p^0(0), q^0(0)),$$

where  $(p^0(0), q^0(0))$  is distributed according to the original equilibrium canonical distribution (associated with (7.1) and  $\beta_N = \beta N^{-1}$ ). Then for all  $t \in [0, T]$  one has the uniform convergence

$$\lim_{\bar{\nu} \rightarrow 0} \limsup_{N \rightarrow +\infty} \mathbb{E} \left[ \|q^{\bar{\nu}}(t) - q^{\bar{\nu}=0}(t)\|_{\ell_2}^2 \right] = 0.$$

*Proof.* We write  $X = (q, p)$ , and introduce the norm

$$\|X\|_{\bar{\nu}} = \|q\|_{\ell_2} + \|M_{\nu_N}^{-1/2} p\|_{h_{-1}}.$$

The system (7.3) becomes a stochastic differential equation in the form

$$dX_t^{\bar{\nu}} = A_{\bar{\nu}} X_t^{\bar{\nu}} + F(X_t^{\bar{\nu}}) + \Sigma dW_t, \quad (7.4)$$

where by definition

$$A_{\bar{\nu}} = \begin{pmatrix} 0 & (\text{Id} - \bar{\nu}^2 \Delta_d)^{-1} \\ \Delta_d & 0 \end{pmatrix}, \quad F(X) = \begin{pmatrix} 0 \\ -v'_{\text{ext}}(q) - \gamma p \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 0 \\ \sqrt{N} \sigma \end{pmatrix}.$$

Duhamel formula gives an implicit expression for differences of solutions of (7.4) with the same noise

$$\begin{aligned} X_t^{\bar{\nu}} - X_t^0 &= (e^{A_{\bar{\nu}} t} - e^{A_0 t}) X_0^0 + \int_0^t (e^{A_{\bar{\nu}}(t-s)} - e^{A_0(t-s)}) (F(X_s^0) ds + \Sigma dW_s) \\ &\quad + e^{t A_{\bar{\nu}}} (X_0^{\bar{\nu}} - X_0^0) + \int_0^t e^{A_{\bar{\nu}}(t-s)} (F(X_s^{\bar{\nu}}) - F(X_s^0)) ds. \end{aligned} \quad (7.5)$$

We estimate the individual terms on the right hand side in (7.5). We define  $P$  as the coordinate transformation associated with the orthonormal spectral decomposition

$$-\Delta_d = P^{-1} \text{diag}(\delta_0, \dots, \delta_{N-1}) P,$$

where  $PP^T = \text{Id}$ . The eigenvalues of the discrete Neumann Laplacian are given, for  $k = 0, \dots, N-1$ , by

$$\delta_k = 4N^2 \sin^2 \left( \frac{k\pi}{2N} \right) \underset{N \rightarrow \infty}{\sim} k^2 \pi^2. \quad (7.6)$$

Denoting the spectral coordinates

$$\hat{X} = (\hat{q}, \hat{p}) = (N^{-1/2} P p, N^{-1/2} P q)$$

we have

$$\|X\|_{\bar{\nu}}^2 = \hat{p}_0^2 + \sum_{k=1}^{N-1} \frac{\delta_k}{1 + \bar{\nu}^2 \delta_k} \hat{p}_k^2 + \sum_{k=0}^{N-1} \hat{q}_k^2.$$

The spectral decomposition leads to a block diagonal form of the operator  $e^{A_{\bar{\nu}} t}$  with diagonal  $2 \times 2$  blocks in the spectral basis

$$e^{\widehat{A_{\bar{\nu}}}(0)t} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},$$

as well as for  $k = 1, \dots, N-1$

$$e^{\widehat{A}_{\bar{\nu}}(k)} = \begin{pmatrix} 0 & (1 + \bar{\nu}^2 \delta_k)^{-1} \\ \delta_k & 0 \end{pmatrix},$$

where  $\widehat{A}_{\bar{\nu}}(k)$  is the  $2 \times 2$  block associated with the coordinates  $(\hat{q}_k, \hat{p}_k)$ . Since  $\widehat{A}_{\bar{\nu}}(k)$  conserves the  $k$ -mode energy  $\delta_k(\hat{q}_k)^2 + (1 + \bar{\nu}^2 \delta_k)^{-1}(\hat{p}_k)^2$ , one can check that for any  $N \geq 1$  the operator norm

$$\|e^{A_{\bar{\nu}} t}\|_{\bar{\nu}}^2 \leq 2 + 2t^2.$$

Similarly, since in the sense of symmetric matrices  $M_{\nu_N}^{-1/2} \leq \text{Id}$ , we have the bound

$$\|F(X^{\bar{\nu}}) - F(X^0)\|_{\bar{\nu}} \leq \left\| M_{\nu_N}^{-1/2} (v'_{\text{ext}}(q^{\bar{\nu}}) - v'_{\text{ext}}(q^0) + \gamma(p^{\bar{\nu}} - p^0)) \right\|_{h_{-1}} \leq (L_F + \gamma) \|X^{\bar{\nu}} - X^0\|_{\bar{\nu}}.$$

Using independence of Brownian increments we compute

$$\mathbb{E} \left[ \left\| \int_0^t (e^{A_{\bar{\nu}}(t-s)} - e^{A_0(t-s)}) \Sigma dW_s \right\|_{\bar{\nu}}^2 \right] = \int_0^t \sum_{i=1}^N \left\| (e^{A_{\bar{\nu}}(t-s)} - e^{A_0(t-s)}) \Sigma_{.,i} \right\|_{\bar{\nu}}^2 ds.$$

Applying Gronwall lemma in (7.5) and collecting all terms we obtain

$$\mathbb{E} \left[ \|X_t^{\bar{\nu}} - X_t^0\|_{\bar{\nu}}^2 \right] \leq C_T \left( \mathbb{E} \left[ \|X_0^{\bar{\nu}} - X_0^0\|_{\bar{\nu}}^2 \right] + m_T \right), \quad (7.7)$$

where  $C_T$  is independent of  $N$ , and with  $X^0$  being distributed canonically  $m_T$  is given

$$m_T = \sup_{t \in [0, T]} \left( \mathbb{E} \left[ \left\| (e^{A_{\bar{\nu}} t} - e^{A_0 t}) X^0 \right\|_{\bar{\nu}}^2 \right] + \mathbb{E} \left[ \left\| (e^{A_{\bar{\nu}} t} - e^{A_0 t}) F(X^0) \right\|_{\bar{\nu}}^2 \right] + \sum_{i=1}^N \left\| (e^{A_{\bar{\nu}} t} - e^{A_0 t}) \Sigma_{.,i} \right\|_{\bar{\nu}}^2 \right).$$

For a given random vector  $X$  such that  $\mathbb{E} \left[ \|X\|_0^2 \right] < +\infty$ , Parseval identity and the inequality  $\|\cdot\|_{\bar{\nu}} \leq \|\cdot\|_0$  imply

$$\mathbb{E} \left[ \left\| (e^{A_{\bar{\nu}} t} - e^{A_0 t}) X \right\|_{\bar{\nu}}^2 \right] = \sum_{k=1}^{N-1} \mathbb{E} \left[ \left\| (e^{\widehat{A}_{\bar{\nu}} t} - e^{\widehat{A}_0 t}) \hat{X} \right\|_{k, \bar{\nu}}^2 \right] \leq 2 \sum_{k=1}^{N-1} \mathbb{E} \left[ \left\| \hat{X} \right\|_{k, 0}^2 \right], \quad (7.8)$$

where  $\|\cdot\|_{k, \bar{\nu}}$  is the restriction to the  $k$ -th mode  $(\hat{q}_k, \hat{p}_k)$ . Then one has, by orthogonality of  $P$ ,

$$\sum_{i=1}^N \mathbb{E} \left[ \left\| \widehat{\Sigma}_{.,i} \right\|_{k, 0}^2 \right] = \sigma^2 \sum_{i=1}^N P_{k,i}^2 \frac{1}{\delta_k} \leq \frac{\sigma^2}{\delta_k}.$$

Up to normalization, the distribution of  $X^0$  has the density  $e^{-\frac{\beta}{N} \sum_{i=1}^d v_{\text{ext}}(q_i^0)}$  with respect to the Gaussian distribution with the covariance matrix  $\beta^{-1} \text{Id}$  for momenta variables, and the covariance matrix  $(\beta \Delta_d)^{-1}$  for positions. Thus we have the bound

$$\mathbb{E} \left[ \left\| \widehat{X}^0 \right\|_{k, 0}^2 \right] \leq 2e^{4\beta \|v_{\text{ext}}\|_{\infty}} \frac{1}{\delta_k \beta},$$

as well as

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[ \left\| \widehat{F(X^0)} \right\|_{k, 0}^2 \right] \leq e^{4\beta \|v_{\text{ext}}\|_{\infty}} \mathbb{E} \left[ \left\| \mathcal{F} \circ F \circ \mathcal{F}^{-1}(\widehat{G}^0) \right\|_{k, 0}^2 \right],$$

where  $\mathcal{F}$  denotes the Fourier series expansion on  $[0, 1]$  with Neumann conditions, and  $(\widehat{G^0}_k)_{k \geq 1}$  are canonical centered Gaussian i.i.d. variables with the covariance matrix  $\beta^{-1} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{k^2 \pi^2} \end{pmatrix}$ . By the Lipschitz assumption the series is bounded

$$\sum_{k=1}^{+\infty} \mathbb{E} \left[ \left\| \mathcal{F} \circ F \circ \mathcal{F}^{-1}(\widehat{G^0}) \right\|_{k,0}^2 \right] \leq (L_v + \gamma) \mathbb{E} \left[ \left\| \widehat{G^0} \right\|_0^2 \right] = \sum_{k=1}^{+\infty} \frac{2(L_v + \gamma)}{\beta k^2 \pi^2}.$$

Since  $\lim_{\nu \rightarrow 0} \left\| e^{\widehat{A_\nu}^{(k)} t} - e^{\widehat{A_0}^{(k)} t} \right\| = 0$ , one can take the limit  $N \rightarrow +\infty$  and use the uniform convergence of the series in (7.8) to obtain  $\lim_{\bar{\nu} \rightarrow 0} \lim_{N \rightarrow +\infty} m_T = 0$  in (7.7). The convergence of the initial condition  $\lim_{\bar{\nu} \rightarrow 0} \lim_{N \rightarrow +\infty} \mathbb{E} \|X_0^{\bar{\nu}} - X_0^0\|_{\bar{\nu}}^2$  follows by using similar arguments. The proof is complete.  $\square$

**7.2. Relaxation of time-step stability restriction.** To demonstrate improved stability properties of time integration algorithms we consider the IMMP scheme (4.1) associated with the mass-matrix penalized Hamiltonian (7.2). Note that when the constraints are linear, the leapfrog scheme (RATTLE) applied to an implicit Hamiltonian is identical to the usual leapfrog scheme for the associated explicit Hamiltonian (7.2). We restrict the rigorous analysis to the quadratic interaction potential ( $v_{\text{int}}(r) = \frac{r^2}{2}$ ), zero exterior potential ( $v_{\text{ext}} = 0$ ), and to the mass-matrix penalization operator ( $\text{Id} - \bar{\nu}^2 \Delta_d$ ). The leapfrog scheme is defined as

$$\begin{cases} p_{n+1/2} = p_n + \frac{\delta t}{2}(-\Delta_d)q_n \\ q_{n+1} = q_n + \delta t M_{\nu_N}^{-1} p_{n+1} \\ p_{n+1} = p_{n+1/2} + \frac{\delta t}{2}(-\Delta_d)q_{n+1}. \end{cases}$$

Denoting the spectral variables for  $k = 1, \dots, N-1$

$$\begin{cases} \widehat{v}^k = \left( \frac{\delta_k}{1 + \bar{\nu}^2 \delta_k} \right)^{1/2} \sqrt{N} P p \\ \widehat{x}^k = \left( \frac{1 + \bar{\nu}^2 \delta_k}{\delta_k} \right)^{1/2} \sqrt{N} P q, \end{cases} \quad (7.9)$$

we write

$$\begin{pmatrix} \widehat{v}_{n+1}^k \\ \widehat{x}_{n+1}^k \end{pmatrix} = L_k \begin{pmatrix} \widehat{v}_n^k \\ \widehat{x}_n^k \end{pmatrix},$$

where

$$L_k = \begin{pmatrix} 1 - \frac{h_k^2}{2} & -h_k + \frac{h_k^3}{4} \\ h_k & 1 - \frac{h_k^2}{2} \end{pmatrix}, \quad \text{and} \quad h_k = \delta t \frac{\delta_k^{1/2}}{(1 + \bar{\nu}^2 \delta_k)^{1/2}}.$$

Since  $\det(L_k) = 1$ , the standard CFL stability condition is equivalent to

$$|\text{Tr}(L_k)| \leq 2$$

which is fulfilled if and only if  $h_k \leq 2$  for all  $k \leq N-1$ . Thus we arrive at the following bound on the time step

$$\delta t \leq 2 \min_{0 \leq k < N} \left( \frac{1 + \bar{\nu}^2 \delta_k}{\delta_k} \right)^{1/2}.$$

Summarizing the above calculations and recalling (7.6) we have the following characterization of the stability properties.

**PROPOSITION 7.2.** *Suppose  $v_{\text{ext}} = 0$  and consider a harmonic interaction potential  $v_{\text{int}}(r) = \frac{r^2}{2}$  with the mass-matrix penalization  $M_{\nu_N} = \text{Id} - \bar{\nu}^2 \Delta_d$ . The leapfrog/Verlet integration of the IMMP harmonic Hamiltonian (7.1) is stable in the spectral sense if and only if*

$$\delta t \leq \left( 4\bar{\nu}^2 + \frac{1}{N^2 \sin^2 \left( \frac{(N-1)\pi}{2N} \right)} \right)^{1/2}. \quad (7.10)$$

Since we work with a Metropolis correction of the hybrid Monte-Carlo type, we are also interested in the limiting behavior of the energy variation compared to the temperature, i.e.,

$$\beta_N (H(p_{n+1}, q_{n+1}) - H(p_n, q_n)),$$

when  $(p_n, q_n)$  are distributed according to the canonical distribution. This quantity gives the average acceptance rate of the Metropolis correction. The result we present here is similar to [4] where the authors analyze infinite dimensional sampling with the standard Metropolis-Hastings Markov chains.

**PROPOSITION 7.3.** *Suppose  $v_{\text{ext}} = 0$  and consider a harmonic interaction potential  $v_{\text{int}}(r) = \frac{r^2}{2}$  with the mass-matrix penalization  $M_{\nu_N} = \text{Id} - \bar{\nu}^2 \Delta_d$ . Suppose the state variable  $X = (p_{\nu_N}, q)$  is a random variable distributed according to the canonical distribution associated with the mass-matrix penalized Hamiltonian (7.1). Then the energy variation  $\beta_N \Delta H$  after one step of the leapfrog integration scheme converges in distribution, up to normalization and centering, to the Gaussian random variable*

$$\frac{\beta_N \Delta H - m_N}{\sigma_N} \xrightarrow[N \rightarrow +\infty]{\text{Law}} \mathcal{N}(0, 1),$$

with the mean and variance in the infinite size asymptotics for the IMMP method  $\bar{\nu} > 0$  and  $\delta t \equiv \delta t_N = o(1)$

$$m_N \underset{N \rightarrow +\infty}{\sim} \frac{N \delta t_N^6}{32 \bar{\nu}^6}, \quad \text{and} \quad \sigma_N^2 \underset{N \rightarrow +\infty}{\sim} \frac{N \delta t_N^6}{16 \bar{\nu}^6},$$

and for the Verlet integration of exact dynamics with  $\delta t \equiv \delta t_N = o(1/N)$

$$m_N \underset{N \rightarrow +\infty}{\sim} \frac{5}{8} N^7 \delta t_N^6, \quad \text{and} \quad \sigma_N^2 \underset{N \rightarrow +\infty}{\sim} \frac{5}{4} N^7 \delta t_N^6.$$

*Proof.* We start with a canonically distributed state  $X = (q, p)$ , which is, by assumption on the form of the interaction potential, a Gaussian random vector. After changing to the spectral coordinates (7.9) we have the spectral representation of the Hamiltonian

$$\beta_N H = \beta \sum_{k=1}^{N-1} \frac{\delta_k^{1/2}}{2(1 + \bar{\nu}^2 \delta_k)^{1/2}} ((\hat{v}^k)^2 + (\hat{x}^k)^2),$$

and introducing Gaussian random vectors  $U$  and  $V$  with the identity covariance matrix we can write

$$\hat{x}^k = \beta^{-1/2} \frac{(1 + \bar{\nu}^2 \delta_k)^{1/4}}{\delta_k^{1/4}} U_k, \quad \text{and} \quad \hat{v}^k = \beta^{-1/2} \frac{(1 + \bar{\nu}^2 \delta_k)^{1/4}}{\delta_k^{1/4}} V_k.$$

We then compute explicitly the change of the Hamiltonian after one step of the leapfrog integration

$$\beta_N \Delta H = \sum_{k=1}^{N-1} \frac{1}{2} \begin{pmatrix} U_k \\ V_k \end{pmatrix}^T (L_k^T L_k - \text{Id}) \begin{pmatrix} U_k \\ V_k \end{pmatrix}. \quad (7.11)$$

Since  $\det(L_k^T L_k) = 1$  the matrix  $L_k^T L_k - \text{Id}$  has two positive eigenvalues  $(\lambda_k - 1, 1/\lambda_k - 1)$  which satisfy

$$\begin{aligned}\lambda_k + 1/\lambda_k - 2 &= \text{Tr}(L_k^T L_k - \text{Id}) = \frac{h_k^6}{16}, \\ (\lambda_k - 1)^2 + (1/\lambda_k - 1)^2 &= \text{Tr}(L_k^T L_k)^2 - 2\text{Tr}(L_k^T L_k) = \frac{h_k^{12}}{256} + \frac{h_k^6}{8}.\end{aligned}$$

Combining with (7.11) we find

$$m_N \equiv \mathbb{E}[\beta_N \Delta H] = \sum_{k=1}^{N-1} \frac{h_k^6}{2^5}, \quad \text{and} \quad \sigma_N^2 \equiv \text{Var}[\beta_N \Delta H] = \sum_{k=1}^{N-1} \frac{h_k^6}{2^4} + \frac{h_k^{12}}{2^9}.$$

Moreover, the Lindeberg or simply Lyapunov condition in the general central limit theorem (see [18]) is verified since we work with a sum of  $\chi^2$  random variables, thus concluding the first part of the proof.

Recalling

$$h_k = \delta t \frac{\sin(\frac{k\pi}{N/2})}{(\frac{1}{4N^2} + \bar{\nu}^2 \sin^2(\frac{k\pi}{N/2}))^{1/2}},$$

we compute the convergent Riemann sums for  $p = 6$  and  $p = 12$ . For the case  $\bar{\nu} \neq 0$  we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^{N-1} h_k^p = \frac{\delta t^p}{\bar{\nu}^p}.$$

If  $\bar{\nu} = 0$  we obtain

$$\begin{aligned}\lim_{N \rightarrow \infty} \frac{1}{N^{p+1}} \sum_{k=1}^{N-1} h_k^p &= \lim_{N \rightarrow \infty} \frac{\delta t^p}{N} \sum_{k=1}^{N-1} 2^p \sin^p \left( \frac{k}{2N} \pi \right) \\ &= \delta t^p 2^p \int_0^1 \sin^p \left( \frac{\pi}{2} x \right) dx.\end{aligned}$$

Thus for  $p = 6$  we have that the series sums to  $20\delta t^6$ . Then the asymptotic behavior follows from the assumption  $\delta t_N^{12} \ll \delta t_N^6$ , and similarly  $\delta t_N^{12} N^{12} \ll \delta t_N^6 N^6$  in the case  $\bar{\nu} = 0$ .  $\square$

**REMARK 7.4.** The two propositions proved in this section characterize the restrictions imposed by the stability of the resulting scheme. In Proposition 7.2, stability in the large system size limit,  $N \rightarrow +\infty$ , is equivalent to the inequality (7.10). In this case the restriction of the time-step size is imposed by the numerical integrator. On the other hand the stability for the scheme which uses a Metropolis corrector is linked to the acceptance rate of the Metropolis step. In Proposition 7.2, stability in the large system size limit is equivalent to the non-vanishing Metropolis acceptance rate, which is equivalent to bounded from above average energy variation  $m_N$  and bounded variance  $\sigma_N$  of the energy variation. In either case, the IMMP method ( $\bar{\nu} > 0$ ) induces a relative increase of order  $N$  for the boundary of numerical stability as compared to the exact dynamics  $\bar{\nu} = 0$  integrated with the Verlet scheme.

**Appendix A. Surface measures.** Let  $\mathbb{R}^d$  be endowed with the scalar product given by the positive definite matrix  $M$ , and consider  $\mathcal{M}_z$  a family of sub-manifolds of co-dimension  $n$  implicitly defined by  $n$  independent functions  $\mathcal{M}_z = \{x \in \mathcal{M} \mid \xi_1(q) = z_1, \dots, \xi_n(q) = z_n\}$  for  $z$  in a neighborhood of the origin. For each  $z$  in a neighborhood of the origin the *conditional measure*  $\delta_{\xi(q)=z}(dq)$  is a measure on  $\mathcal{M}_z$  defined in such a way that it satisfies the chain rule for conditional expectations with respect to the Lebesgue measure  $dq$ , i.e.,

$$dq = \delta_{\xi(q)=z}(dq) dz. \quad (\text{A.1})$$

The *surface measure*  $\sigma_{T_q^* \mathcal{M}_z}(dp)$  is the Hausdorff measure induced by the metric  $M^{-1}$  on the co-tangent space  $T_q^* \mathcal{M}_z = \{p \mid \nabla_q^T \xi(q) M^{-1} p = 0\}$ ; and in the same way,  $\sigma_{\mathcal{M}_z}(dq)$  is the Hausdorff measure induced by the metric  $M$  on the sub-manifold  $\mathcal{M}_z$ . It is important to note that, although this is not explicit in the notation,  $\sigma$  is defined with respect to the mass-tensor  $M$  of the mechanical system. The Liouville measure  $\sigma_{T^* \mathcal{M}_z}(dp dq)$  on the co-tangent bundle  $T^* \mathcal{M}_z$  is the volume form induced on

$$T^* \mathcal{M}_z = \{(p, q) \mid \nabla_q^T \xi(q) M^{-1} p = 0, \xi(q) = z\}$$

by the usual symplectic form  $dp \wedge dq$ . It can be described in terms of surface measures as follows

$$\sigma_{T^* \mathcal{M}_z}(dp dq) = \sigma_{T_q^* \mathcal{M}_z}(dp) \sigma_{\mathcal{M}_z}(dq).$$

Finally, the co-area formula (see [17] for a general reference) defines the relative probability density between  $\delta_{\xi(q)=z}(dq)$  and  $\sigma_{\mathcal{M}_z}(dq)$ .

**PROPOSITION A.1** (Co-area formula). *Given the invertible Gram matrix associated with the constraints  $\xi(q) = z$  in a neighborhood of  $\mathcal{M}_z = \{q \mid \xi(q) = z\}$*

$$G(q) = \nabla_q^T \xi M^{-1} \nabla_q \xi,$$

one has

$$\delta_{\xi(q)=z}(dq) = \frac{1}{\sqrt{\det G(q)}} \sigma_{\mathcal{M}_z}(dq).$$

## Appendix B. Langevin processes.

Defining the Poisson bracket

$$\{\varphi_1, \varphi_2\} = \nabla_p^T \varphi_1 \nabla_q \varphi_2 - \nabla_p^T \varphi_2 \nabla_q \varphi_1,$$

and the dissipation tensor

$$d(q) = \sigma q,$$

where  $\sigma$  is the fluctuation matrix in Definition 2.1, the Markov generator of the Langevin process in Definition 2.1 is

$$\mathcal{L} = \{\cdot, H\} + \frac{1}{\beta} \{d, \{d^T, \cdot\} e^{-\beta H}\} e^{\beta H}.$$

The generator  $\mathcal{L}$  satisfies

$$\int \varphi_1 \mathcal{L}(\varphi_2) e^{-\beta H} dp dq = \int \mathcal{L}^*(\varphi_1) \varphi_2 e^{-\beta H} dp dq,$$

where

$$\mathcal{L}^* = \{\cdot, -H\} + \frac{1}{\beta} \{d, \{d^T, \cdot\} e^{-\beta H}\} e^{\beta H}.$$

The generator  $\mathcal{L}^*$  defines a Langevin process with the time-reversed Hamiltonian  $(-H)$ . Reversibility of the process implies that the canonical measure is stationary. Furthermore, if the initial state of the system is a canonically distributed random variable, the probability distribution of a trajectory after the time-reversal is given by a Langevin process with the generator  $\mathcal{L}^*$ . When  $H$  has the form  $H(p, q) = \frac{1}{2} p^T M^{-1} p + V(q)$ , reversal of impulses ( $p \rightarrow -p$ ) leads to time-reversed dynamics, and a process with generator  $\mathcal{L}^*$  can be constructed by the following simple steps:

1. Reverse momenta ( $p \rightarrow -p$ ).
2. Draw a random path with generator  $\mathcal{L}$ .
3. Reverse again momenta ( $p \rightarrow -p$ ).

When holonomic constraints, for instance, of the form

$$\Xi(p, q) = \zeta \Leftrightarrow \begin{cases} p^T M^{-1} \nabla_q \xi = 0 \\ \xi(q) = z \end{cases}$$

are introduced, it is useful to define the Poisson bracket on the co-tangent bundle  $T^* \mathcal{M}_z$

$$\{\varphi_1, \varphi_2\}_{\mathcal{M}_z} = \{\varphi_1, \varphi_2\} - \sum_{a,b} \{\varphi_1, \Xi^a\} \Gamma_{a,b}^{-1} \{\Xi^b, \varphi_2\} ,$$

where  $\Gamma$  is the symplectic Gram matrix of the full constraints

$$\Gamma^{a,b} = \{\Xi^a, \Xi^b\} .$$

As a basic result of symplectic geometry (see [1]), one recovers the divergence formula with respect to the bracket  $\{\cdot, \cdot\}_{\mathcal{M}_z}$  and the Liouville measure  $\sigma_{T^* \mathcal{M}_z}(dp dq)$

$$\int \{\cdot, \cdot\}_{\mathcal{M}_z} \sigma_{T^* \mathcal{M}_z}(dp dq) = 0 .$$

Given a constrained Langevin process in a stochastic differential equation form

$$\begin{aligned} \dot{q} &= \nabla_p H , \\ \dot{p} &= -\nabla_q H - \gamma \dot{q} + \sigma \dot{W} - \nabla_q \xi \dot{\lambda} , \end{aligned}$$

where  $\lambda$  are Lagrange multipliers associated with the constraints  $\xi(q) = 0$ , adapted with respect to the noise  $\dot{W}$ , the process  $\{p_t, q_t\}_{t \geq 0}$  obeys hidden velocity constraints and is characterized by the stochastic differential equations

$$\begin{aligned} \dot{q} &= \nabla_p H + \nabla_p \Xi \dot{\Lambda} , \\ \dot{p} &= -\nabla_q H - \gamma \dot{q} + \sigma \dot{W} - \nabla_q \Xi \dot{\Lambda} , \end{aligned}$$

where  $\Lambda$  are Lagrange multipliers associated with the full constraints  $\Xi(p, q) = 0$ . The Markov generator of this process can be written in the form

$$\mathcal{L}_{\mathcal{M}_z} = \{\cdot, H\}_{\mathcal{M}_z} + \frac{1}{\beta} \left\{ d, \{d^T, \cdot\}_{\mathcal{M}_z} e^{-\beta H} \right\}_{\mathcal{M}_z} e^{\beta H} ,$$

demonstrating the reversibility with respect to the constrained canonical measure  $e^{-\beta H} \sigma_{T^* \mathcal{M}_z}(dp dq)$ .

**Appendix C. Exact sampling of fluctuation/dissipation perturbations.** In this section, we recall how to perform exact sampling of fluctuation/dissipation perturbations. Since we only work with impulses, we refer to the system by using the impulse variables  $p$  only. Note that throughout the paper, we also use extended variables  $(p, p_z)$ , however, the presentation that follows covers general cases. The kinetic energy of the system is  $\frac{1}{2} p^T M p$ . We impose constraints  $p^T M^{-1} \nabla_q \xi = 0$  on impulses, thus  $p \in T_q^* \mathcal{M}$  and hence the associated orthogonal projector on  $T_q^* \mathcal{M}$  is

$$P = \text{Id} - \nabla_q \xi G^{-1} \nabla_q^T \xi M^{-1} .$$

The stochastic differential equations of motion on impulses that are integrated on a time-step interval are

$$\begin{cases} \dot{p} = -\gamma M^{-1}p + \sigma \dot{W} - \nabla_q \xi \dot{\lambda}, \\ p^T M^{-1} \nabla_q \xi = 0, \end{cases} \quad (C_p) \quad (C.1)$$

with the usual fluctuation/dissipation relation  $\sigma \sigma^T = 2\beta^{-1}\gamma$ . The Gaussian distribution of impulses

$$\frac{1}{Z} e^{-\frac{\beta}{2} p^T M^{-1} p} \sigma_{T_q^* \mathcal{M}}(dp) \quad (C.2)$$

is invariant under the dynamics (C.1).

**PROPOSITION C.1** (Exact sampling of stochastic perturbation). *Given the mass matrix  $M$ , suppose either  $\delta t$  or  $\gamma$  are small enough so that the condition*

$$\frac{\delta t}{2} M^{-1} \leq \gamma \quad (C.3)$$

*holds in the sense of symmetric semi-definite matrices. Let  $U$  be a centered and normalized Gaussian vector. Consider the mid-point Euler scheme with constraints*

$$\begin{cases} p_{n+1} = p_n - \frac{\delta t}{2} \gamma M^{-1} (p_n + p_{n+1}) + \sqrt{\delta t} \sigma U - \nabla_q \xi \lambda_{n+1} \\ p_{n+1}^T M^{-1} \nabla_q \xi = 0, \end{cases} \quad (C_p) \quad (C.4)$$

*where  $\lambda_{n+1}$  is the Lagrange multiplier associated with the constraint  $(C_p)$ . The Markov kernel defined by the transition  $p_n \rightarrow p_{n+1}$  is reversible with respect to the Gaussian distribution (C.2).*

*Proof.* After calculating the Lagrange multiplier the expression (C.4) can be written as

$$p_{n+1} = p_n - \frac{\delta t}{2} P \gamma P^T M^{-1} (p_n + p_{n+1}) + \sqrt{\delta t} P \sigma U.$$

Consider the new variable  $\tilde{p} = \beta^{1/2} M^{-1/2} p$ , and define the symmetric matrix

$$L \equiv \frac{\delta t}{2} M^{-1/2} P \gamma P^T M^{-1/2},$$

as well as  $K$ , such that  $KK^T = L$ . In terms of these new variables we obtain from (C.4)

$$\tilde{p}_{n+1} = (\text{Id} + L)^{-1} (\text{Id} - L) \tilde{p}_n + 2(\text{Id} + L)^{-1} K U. \quad (C.5)$$

Moreover, the product measure  $\sigma_{T_q^* \mathcal{M}_0}(dp_n) \sigma_{T_q^* \mathcal{M}_0}(dp_{n+1})$  is the measure induced on the linear subspace of constraints by the scalar product  $M^{-1}$  and the Lebesgue measure  $dp_n dp_{n+1}$ . Thus in the variables  $(\tilde{p}_n, \tilde{p}_{n+1})$  this measure becomes, up to a constant, the measure induced by the usual Euclidean structure. As a consequence the log density of the random variable  $(\tilde{p}_n, \tilde{p}_{n+1})$  defined by (C.5) with respect to this latter measure is equal to

$$\begin{aligned} & -\frac{1}{2} |\tilde{p}_n|^2 - \frac{1}{8} (\tilde{p}_{n+1} - (\text{Id} + L)^{-1} (\text{Id} - L) \tilde{p}_n)^T L^{-1} (\text{Id} + L)^2 (\tilde{p}_{n+1} - (\text{Id} + L)^{-1} (\text{Id} - L) \tilde{p}_n) \\ & = -\frac{1}{8} \tilde{p}_{n+1}^T L^{-1} (\text{Id} + L)^2 \tilde{p}_{n+1} - \frac{1}{8} \tilde{p}_n^T L^{-1} (\text{Id} + L)^2 \tilde{p}_n, \end{aligned}$$

which is indeed symmetric between  $\tilde{p}_n$  and  $\tilde{p}_{n+1}$ . Hence we have shown the reversibility of the induced Markov kernel and consequently stationarity of the canonical Gaussian distribution.  $\square$

## REFERENCES

- [1] V. I. Arnol'd. *Mathematical methods of classical mechanics*. Springer-Verlag, New York, 1989.
- [2] M. Beccaria and G. Curci. The Kramers equation simulation algorithm: 1. Operator analysis. *Phys. Rev. D*, 49:2578–2589, 1994.
- [3] C.H. Bennett. Mass tensor molecular dynamics. *J. Comp. Phys.*, 19:267–279, 1975.
- [4] A. Beskos, G.O. Roberts, A.M. Stuart, and J. Voss. An MCMC method for diffusion bridges. *Pre-print*, 2007.
- [5] F. Bornemann and C. Schuette. Homogenization of Hamiltonian system with a strong constraining potential. *Physica D*, 102:57–77, 1992.
- [6] C. Le Bris and F. Legoll. Derivation of symplectic numerical schemes for highly oscillatory Hamiltonian systems. *C. R. Acad. Sci. Paris*, 344:277–282, 2007.
- [7] E. Cancès, F. Legoll, and G. Stoltz. Comparison of NVT sampling methods. Technical Report 2040, IMA, 2005.
- [8] G. Ciccotti, T. Lelièvre, and E. Vanden-Eijnden. Projection of diffusions on submanifolds: Application to mean force computation. Technical Report 309, CERMICS, 2006.
- [9] M. E. Clamp, P. G. Baker, C.J. Stirling, and A. Brass. Hybrid Monte Carlo : an efficient algorithm for condensed matter simulation. *J. Comput. Chem.*, 15(8):838–846, 1994.
- [10] D. Cohen, T. Jahnke, K. Lorenz, and Ch. Lubich. Numerical integrators for highly oscillatory Hamiltonian systems: a review. In Aexander Mielke, editor, *Analysis, Modeling and Simulation of Multiscale Problems*, pages 553–576. Springer, 2006.
- [11] M. Creutz and A. Gocksch. Higher-order hybrid Monte Carlo algorithms. *Phys. Rev. Lett.*, 63(1):9–12, 1988.
- [12] S. Duane. Stochastic quantization versus the microcanonical ensemble - getting the best of both worlds. *Nuclear Physics B*, 257(5):652–662, 1985.
- [13] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth. Hybrid Monte Carlo. *Phys. Lett.*, 195B(2):216–222, 1987.
- [14] C. Lubich E. Hairer, G. Wanner. *Geometric Numerical Integration, Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer, 2002.
- [15] G. Ciccotti E. Vanden-Eijnden. Second-order integrators for Langevin equations with holonomic constraints. *Chem. Phys. Lett.*, 429(1-3):310–316, 2006.
- [16] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics, 1986.
- [17] H. Federer. *Geometric measure theory*. Springer, New-Yoerk, 1969.
- [18] W. Feller. *An Introduction to Probability Theory and Its Applications*. Wiley, New-York, 1971.
- [19] M. Fixman. Simulation of polymer dynamics. I. General theory. *J. Chem. Phys.*, 69:1527, 1978.
- [20] M. Fixman. Implicit algorithm for Brownian dynamics of polymers. *Macromolecules*, 19:1195–1204, 1986.
- [21] E. Vanden-Eijnden G. Ciccotti, R. Kapral. Blue moon sampling, vectorial reaction coordinates, and unbiased constrained dynamics. *J. Chem. Phys.*, 6(9):1809–14, 2005.
- [22] A. M. Horowitz. A generalized guided Monte Carlo algorithm. *Phys. Lett.B*, 268:247–252, 1991.
- [23] I. Horváth and A. D. Kennedy. The Local Hybrid Monte Carlo algorithm for free field theory. *Nucl. Phys. B*, 510:367–400, 1998.
- [24] J. A. Izaguirre and S. S. Hampton. Shadow hybrid Monte Carlo: an efficient propagator in phase space of macromolecules. *J. Comput. Phys.*, 200(2):581–604, 2004.
- [25] H.J.C. Berendsen J-P. Ryckaert; G. Ciccotti. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-Alkanes. *J. of Comp. Phys.*, 23:327–341, 1977.
- [26] H.J.C. Berendsen K. A. Feenstra, B. Hess. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. of Comp. Chemistry*, 20:786–798, 1999.
- [27] N.G. Van Kampen. Elimination of fast variables. *Physics Reports*, 124(2):9–160, 1985.
- [28] B. J. Leimkuhler and R. D. Skeel. Symplectic numerical integrators in constrained Hamiltonian systems. *J. Comput. Phys.*, 112(1):117–125, 1994.
- [29] B. Mao and A.R. Friedman. Molecular dynamics simulation by atomic mass weighting. *Biophysical Journal*, 58:803–805, 1990.
- [30] G. Martin and J.I. Siepmann. Transferable potentials for phase equilibria. I. United-atom description of *n*-alkanes. *J. Phys. Chem.*, 102:2569–2577, 1998.
- [31] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equations of state calculations by fast computing machine. *J. Chem. Phys.*, 21:1087–1091, 1953.
- [32] B. Oksendal. *Stochastic differential equations (3rd ed.): an introduction with applications*. Springer-Verlag, 1992.
- [33] D. C. Rapaport. *The Art of Molecular Dynamics Simulations*. Cambridge University Press, 1995.
- [34] S. Reich. Smoothed dynamics of highly oscillatory Hamiltonian systems. *Physica D*, 89:28–42, 1995.
- [35] S. Reich. Smoothed Langevin dynamics of highly oscillatory systems. *Physica D*, 138:210–224, 2000.
- [36] S. Reich and B. Leimkuhler. *Simulating Hamiltonian Dynamics*, volume 14 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, 2005.
- [37] H. Rubin and P. Ungar. Motion under a strong constraining force. *Comm. Pure Appl. Math*, 10:65–87, 1957.
- [38] Ch. Schütte, A. Fischer, W. Huisings, and P. Deuflhard. A direct approach to conformational dynamics based on hybrid Monte Carlo. *J. Comput. Phys.*, 151(1):146–168, 1999.

- [39] F. Takens. Motion under the influence of a strong constraining force. In *Global theory of dynamical systems (Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1979)*, volume 819 of *Lecture Notes in Math.*, pages 425–445. Springer, Berlin, 1980.
- [40] W.F. van Gunsteren and H. J. C. Berendsen. Algorithms for macromolecular dynamics and constraint dynamics. *Molecular Physics*, 34(5):1311–1327, 1977.
- [41] W.F. van Gunsteren and H.J.C. Berendsen. Algorithms for Brownian dynamics. *Molecular Physics*, 45:637–647, 1982.